OpenGeoportal Metadata Working Group Virtual Meeting
Thursday, May 29, 2014

Marc McGee, notes
recording URL:  http://goo.gl/2A2rUf  [First 2-3 minutes of meeting is not recorded]

Attendance: ~20 members of the Metadata Working Group (MWG) participated in the meeting

Announcements

● Eric Glass announced that Columbia currently has a beta version of OpenGeoportal running: http://culspatial.cul.columbia.edu/  Currently has metadata search capabilities only and will add Web Services for data in the future.

Metadata Harvester & Metadata Toolkit updates

Metadata Harvester [Chris Barnett]
● Tool for scheduled, batch harvesting of metadata from OGP nodes, csw sources, GeoNetwork instances, and web directories.
● The ingest tool, scheduler, reporting portions of the Harvester are complete and the code is available in the OGP GitHub repository. These portions of the Harvester are working and ready for use.
● Chris is currently working on developing the upload metadata files function.

   Action item: Chris will send to the listserv a URL for a demo version of the Harvester.

Metadata Toolkit [Chris Barnett]
● Web-based tool for the rapid creation and editing of OGP-compatible metadata.
● Identified core fields for a minimal record as well as recommended fields for the initial Metadata Toolkit iteration.
● Crosswalks for FGDC and ISO have been identified.
● Specifications for the initial design and information architecture have been outlined.
● Tufts and other members of OGP community met in May with GeoCat, who will be developing the Metadata Toolkit. The group previewed the upcoming GeoNetwork release; the Metadata Toolkit will be based on root technologies of GeoNetwork.
● The group clarified use cases. It is intended to be a tool both for researchers and student workers. 3 basic metadata forms: minimal, intermediate (core recommended fields), and full editor (may not make it into 1st release).
● Core fields need to be pared down to more minimal set especially for the research. The group also discussed mechanisms for engaging the researcher to provide more thorough metadata (e.g. social media techniques).
● Development to start soon with delivery expected late Summer/early Fall 2014.
● The tool will also include the ability: to search OGP metadata for existing metadata records, to create and store local templates, and to access built-in thesauri and codeset lists.
Chris is available to answer questions about the Harvester or Toolkit development.

Metadata Toolkit decision points [Marc McGee]

In the discussions with GeoCat regarding Metadata Toolkit development, four elements were identified as needing further clarifications as far as best practices and standardization of what elements to use to record information as well as how that information should be recorded. Background on these decision points are outlined in the document: [http://goo.gl/c9Vfdz](http://goo.gl/c9Vfdz)

- **Host Institution** - It was agreed that the best approach for recording host institution will be to use the distribution information elements in FGDC/ISO. The first iteration of distributor ‘contact organization’ is to be used for the OGP host institution. It is best to record the name of the institution in a consistent manner. It is possible we may want to build a list of variations in institutional name possibilities and have a mapping to the controlled version of the host institution name. Also we may want to have a label name for a shorter display version of the host institution name.

  **Action item:** Marc will update the Best Practices documents to reflect this decision.

- **Web Services** - There are many components that go into encoding web services information in metadata records; it was proposed that a Task Group would be the best way to address identifying best practices for this.

  **Action item:** A Task Group identifying best practices for recording web services metadata will include: Chris Barnett, Keith Jenkins, Kim Durante, Susan Powell, Eric Glass, Marc McGee, and Kevin Dyke. Marc will initiate work of the Task Group by sending out the current status to the Task Group volunteers.

- **Data Type** - Top level facets used for searching the type of data in OGP. Current interface selections are: point, line, polygon, raster, scanned map, and library record [MIT]. What other options might our users want for search faceting? Populating these data types from the metadata record requires a complex parsing algorithms against multiple metadata elements, is there a more consistent way to explicitly record data type information in the metadata record to make the parsing easier?

  **Action item:** A Task Group consisting of Jennie Murack, Keith Jenkins, Eric Glass, and Marc McGee in consultation with Chris Barnett will survey the OGP community, identify top-level data types for OGP, identify which data types can be reliably mapped from the current FGDC/ISO values, propose potential additions to the codeset, and propose Best Practices for recording this information in the metadata.

- **Collections** - What is the best way in the metadata to identify a layer that is part of a larger collection of layers in OGP? Proposed approaches include using the series information or the
larger work citation information or cross references information.

**Action item:** Susan Powell and Marc McGee will form a Task Group to address best practices for recording collection information.

**Governance Task Group update** [Marc McGee]

MWG Governance: [http://goo.gl/qekQUq](http://goo.gl/qekQUq)

Jaime Stoltenberg, Ryan Mattke, Alicia Morris and Marc McGee revised the MWG Governance document to reflect comments made at last MWG Virtual Meeting and other comments made by OGP community members. The Governance Task Group is recommending the current draft be officially adopted by the MWG Voting Members.

**Action item:** Marc will send out an update on the status MWG Governance to the entire OGP list and put out a call for a vote of the MWG Voting Members to adopt the current draft as a working Governance document. MWG Voting Members may propose changes to the Governance document at any time.

**Best Practices Task Group update** [Jennie Murack]

Progress has been made by the Best Practices and Thesauri Task groups on the General Best Practices documentation [http://goo.gl/UtvUfP](http://goo.gl/UtvUfP). There are still some outstanding decisions [commented on in this draft] to be made as a group before we can begin to address FGDC and ISO best practices documents. As a group we went through the list and discussed possible approaches to outstanding items.

- Spatial data set specific elements save for standard-specific documents
- Include an introduction
- Title - include punctuation between elements, spelling out place names, titles with established titles original keep title [supplement extra title elements when necessary], order of title elements recommended with understanding order may not be consistent.
- Thesaurus for publishers? Recommended not required
- Outline date format in schema-specific documents
- Host institution names from a list with alias with shorter versions

**Action item:** Jennie Murack will incorporate feedback from the Group comments into the General Best Practices documentation. The updated document will be sent around for final approval. Once this document is complete the MWG will begin to work on standard-specific documentation.

**GeoBlacklight update** [Kim Durante]

An announcement went out to the OGP listserv about the GeoBlacklight project; Stanford now
has a developer, Jack Reed, working on a web mapping interface and they are beginning to build some design documents with feature requirements.

**Action item:** Kim will send out via the listserv information she is able to share about the current status of the GeoBlacklight project.

**Stanford Place Name Linking Work [Kim Durante]**

Kim gave an update of the work that Stanford is doing with the linking of place names between GeoNames and LCSH Geographic Names. This mapping allows Stanford to include GeoNames keywords in ISO records while automatically converting those keywords to LCSH when converting from ISO to MODS records. MODS and ISO both have the ability to include URIs in xlinks as an attribute for textual place names to uniquely identify the place name as well making the keyword data more machine processable.

**Action item:** Kim will send out via the listserv examples of the tables and the work Stanford’s doing on place name linking. Comments and contributions are welcome.

**Schedule next Virtual Meeting**

- The next OGP MWG Virtual Meeting will be held in late August or early September.
- A scheduling poll will go out to the Group via listserv in July and a call for topics will go out prior to the meeting.