19 Mar 2015

Time: 1pm Eastern / 10am Pacific.

Call-In Info: Dial 1-530-881-1400, room 651025#

Moderator: Darren Hardy (Stanford)

Notetaker: Eliot Jordan (Princeton)

Attendees:
- Bess Sadler
- Darren Hardy
- Jack Reed
- Kim Durante
- Tom Britnacher
- Lynne Grigsby
- John Huck
- James R. Griffin III
- Eliot Jordan
- Matthew Sisk
- Jay Varner
- Robin Shane Coleman
- Ed Brooks (Virginia Tech)
- Wayne Graham (U. of Virginia)

Agenda:
1. Roll Call
2. Call for Agenda Items
3. Next Call
   a. Date: 16 Apr 2015
   b. Moderator: Jack Reed
   c. Notetaker: Bess Sadler
4. Review communication channels (hydra-community/hydra-tech mailing lists)
   a. IRC Channels
      i. General recommendation is to use the main hydra channel until traffic increases to point where a new channel is needed
      ii. Server: chat.freenode.net
      iii. Channel: #projecthydra
5. Discussion on institutions’ next steps toward a geospatial Hydra implementation
   a. Stanford
      i. Launching of Hydra implementation for spatial data in April
      ii. Working with PR department and timing launch to coincide with Earth Day
         1. Need to prove that the library is part of the scientific data management workflow and is relevant to the sciences to keep funding levels up
      iii. Culmination of several years of work accessioning GIS data and establishing workflows for:
         1. Metadata creation
         2. Loading datasets into a spatial data infrastructure
         3. Adding layers into a discovery service
      iv. GIS data is modeled in the repository for long term preservation
      v. Using common hydra ecosystem components: Fedora, Solr, Rails, Blacklight, GeoBlacklight, Solr
      vi. Looking to migrate from Fedora 3 to 4 with linked data features and model GIS data as linked data
      vii. Next steps are the launch and accessioning more GIS content
      viii. Gearing up to do a GIS self-deposit interface within the next year
           1. Like Sufia for shapefiles
           2. Interested in collaborators
      ix. One challenge is scanned maps without georeferenced control points
          1. Looking at how to accession georeferenced content into the repository that relates it back to the original scan and paper map
          2. Potentially using crowd sourcing for the georeferencing work
   b. Berkeley
      i. In transition, but the plan is to go down the path with hydra
      ii. Self deposit functionality on hold
      iii. Currently have an OpenGeoPortal instance
      iv. Have many digitized maps, and looking into how best to get metadata and spatial information for these
   c. Univ. of Alberta
      i. The goal is to use Hydra as the repository for all digital collections
      ii. Starting with implementing Sufia as an institutional repository in the next couple of months
      iii. Historical collections, including maps, are next
      iv. Scanned maps and shapefiles will probably come in as parts of other collections
      v. Scanned maps will be first follow by historical census data
      vi. Alberta does have a Fedora repository with a custom interface, but are looking to migrate to Fedora 4
   d. UC Santa Barbara
      i. Starting over on hydra project with new fedora, but haven’t gotten to geospatial.
      ii. Currently housing electronic dissertations and thesis, but will be moving on to special collections and scanned images
      iii. Waiting for the geo component of the project to begin, but are looking at to using GeoBlacklight for discovery
iv. Right now, focusing on description and developing metadata standards and workflows for getting materials ready to go  
v. Have a lot scanned maps with MARC records with coordinate information, and are looking for ways to extract information from those as a starting point  

e. Princeton  
i. Starting a twelve week push to implement hydra for a number a number of resource types including digitized books and born-digital materials  
ii. Geospatial will not be part of these initial sprints, but scanned maps and geo datasets are on the horizon  
iii. Plan to begin testing GeoBlacklight instance this summer for a potential roll-out in August  

f. Notre Dame  
i. Hydra repository is one month old, and are in the planning stages for anything geospatial  
ii. Repository is set up almost entirely for self-deposit  
iii. Looking forward to self-deposit GIS material, but unsure of the demand  
iv. Some of the geohydra tools are already rolled out on the smaller grant funded repository VecNet.  

g. Univ. of Virginia  
i. Migrating GIS content from a 10 year-old custom system  
ii. Skipping the hydra component for now, but are in contact with another group in charge of hydra implementation at the university  
iii. Hope to have a production service by the beginning of the fall semester  

h. Lafayette College  
i. Heavily invested in the Islandora system for digital repository front end  
ii. Working on two early prototypes for managing geospatial datasets  
iii. Managing shapefiles with a content model somewhat similar to what one finds in the Islandora community, and trying to refactor that into a community solution  
iv. Converting shapefiles into KML and GeoJSON, and simplifying the geometry with Mike Bostock's TopoJSON tool.  

v. Exploring the development of a Sufia instance in the Spring semester  
vi. Have issues related to georeferenced tiffs.  
1. Currently managing as Islandora large image objects  
2. Have had some success serving these as raster maps via geoserver  

i. Emory  
i. In the early stages of adopting hydra  
ii. Attended most recent HydraCamp  
iii. Have Geoserver and an instance of OpenGeoPortal  
iv. Looking to migrating to Fedora 4  

j. Virginia Tech  
i. Development group working on Fedora/Hydra repository  
ii. Working with GIS librarian on test instance of GeoBlacklight  
iii. University has a number of groups using GIS; some using GeoServer and others using ESRI products  
iv. Each department wants to maintain their own data  
1. Idea is to standardize metadata and host it through the library's GeoBlacklight instance  
2. One discovery system points to data stored in each department  

6. Hydra resources for the Interest Group  

a. Conferences, meetings, workshops  
   i. Hydra Camp  
   ii. Open Repositories: Jun 8-11, Indianapolis  
   iii. Hydra Connect: Fall 2015, Minneapolis  

b. Identify specific common issue(s) for WG, perhaps:  
   a. XSLT to transfer between metadata formats  
      1. Stanford has a collection of XSLTs for converting ISO and FGDC to MODS and GeoBlacklight Schema  
      2. **Kim Durante will make these available to the group**  
      3. Jack commented that, in general, there isn't a great method for doing this and suggested that GeoCombine might be a good platform for developing these tools  
      4. Kim mentioned that there are style sheets to convert from FGDC to XSLT  
         a. Certain elements, IDs for instance, require extra work on the part of the institution  
         b. This conversion is complicated by the fact that ISO standards have changed  
      5. Darren spoke about GeoBlacklight Schema  
         a. Dublin Core schema with additional elements that provide bounding box and other geospecific metadata  
         b. Provides links to web services such as WMS  
         c. Schema can be visualized with standard DC tools as well as ingested into GeoBlacklight  
   
   b. Issues at the planning stage for moving forward  
      i. Migrating from Fedora 3 to Fedora 4  
         1. John mentioned that they are working to implement Sufia with Fedora 4 and that from a metadata point of view, this represents a lot of change  
         2. Perhaps group can look at the implications of Fedora 4 and RDF?  
         3. There are a lot of unanswered questions about how Fedora 4 handles metadata that is more complex than basic Dublin Core  
         4. Lafayette college is looking to explore this in December  
      ii. Querying against Linked Data  
         1. James asked if there are there any services built on top of or integrated into Fedora 4 that would support querying similar to SPARQL  
         2. Jack and Darren thought is was a good idea to split the administrative layer, which might use tools like SPARQL, from the end-user discovery layer into separate concerns  
   
   c. Shared Data Model  
      i. The group decided that there was enough interest in developing a shared model for GIS data in Fedora 4 to initiate a working group  
      ii. Initial participants include: Darren and Bess from Stanford, James from Lafayette College, John from Alberta, Eliot from Princeton, and Jay from Emory.  
      iii. Bess will reach out to contact at the University of New Hampshire who is currently working on this  
      iv. **Eliot Jordan volunteered to be facilitator**