Development Call 20121018

Calls are held every Thursday at 1 pm eastern daylight time (GMT-5) – convert to your time at http://www.thetimezoneconverter.com

These calls now use WebEx – see the "Call-in Information" at the bottom of this page.

*Please add additional agenda items or updates – *

Jon Corson-Rikert will be leading this week's call.

VIVO 1.5.1

VIVO 1.5.1 was released Monday, October 15. Many thanks to all who contributed through bug reporting, development, and testing.

Updates

- Brown - *have been loading all the data into VIVO and building fairly detailed faculty profiles; hoping to have faculty starting to look in the next month. Have quite a bit of data, all off the CVs that faculty members are asked to produce based on patterns established by the University. Faculty members or their delegates will be asked to update in VIVO after this initial load. The main Brown website is running on Drupal and departments are encouraged to use Drupal as well; no discussions yet on feeding websites from VIVO but that's certainly a goal.*
- Colorado - *Stephen has started and is looking at the upgrade path from 1.2 to 1.5.1. Since they don't allow editing it may be easiest to modify their input scripts to produce 1.5.1 data; they would otherwise need to go through each intervening major release (1.3, 1.4.1, and 1.5.1).*
- Cornell - *working with multiple named graphs to allow adding/removing subsets of data as required.*
- Duke
- EPA
- Florida – *preparing to upgrade to 1.5.1*
- Indiana – *still working with the Florida data to test the scale of Maps of Science that can be drawn. Recommends using the* Jabbr service *from Cornell to align journal names with those used by the base map of science in that visualization (an amalgamation of journals from Web of Science and Scopus)*
- NYU - *tweaking the interface to be optimized for searching and displaying people*
- Penn -
- UCSF
- Weill - Continued to get help from Stella on identifying "unlinked" data. Next step - look into the data and determine what to delete.

```
select
(count (distinct ?o) as ?count)
where {
graph <http://vitro.mannlib.cornell.edu/default/vitro-kb-2> {
?s ?p ?o .
filter (not exists { ?o ?p2 ?o2})
}
}
```

This query will be returning all the ref:type statements and all data properties – which likely explains the large number of values returned (273,000 +) and the time it took to run.

Demo of RIS2N3 Tool

David Cliffe, Indiana University:

So I'm looking to make available a tool I developed that helped with a VIVO implementation here at IU – I've had a look at the wiki, and found a link to VIVO tools https://sourceforge.net/p/vivo/tools/home/Home/

That page provides a link to contact to nominate additional tools, which is the http://vivoweb.org/contact

I figured it may make more sense to just email the dev list.

The tool in question is RIS2N3 - https://github.com/dgcliff/RIS2N3

It converts RIS plain text files containing citations into N3, and in doing so, also creates the predicate relationships between authors and publications (and journals and articles etc.)

It's not complete, and probably never will be, but I thought it may be of use to others. We had a poster about it at the last VIVO conference, where a few people mentioned interest.

Notes from the presentation:

- Now checks against an existing VIVO application by doing a SPARQL query ahead of time to retrieve a list of name parts – checking for people using a simple last name/first initial match that could be improved and applied to article titles, journals and/or publishers
- Wrote his own parser for the RIS format
- At IU they use it to ingest data from Endnote and Google Scholar – RIS is a common export format

- Stella commented that she wrote a similar tool for Scripps that uses a bibtex parser and also checks for matches to existing journals and articles by title; these tools could be merged.
- Stephen mentioned that the Harvester has a workflow designed to take MODS bibliographic metadata as input; an open source library was used to produce MODS metadata from other formats.

Moving to Git/GitHub, Converting the Wiki

We plan to convert the Vitro and VIVO repositories from Subversion to Git, and store them on GitHub. This will likely occur within a week or two of release 1.5.1
Jim is using the SVN to Git tool, which besides straight conversion handles the conversion of the tags that have a different meaning in Git from SVN. He is also restructuring the repository so it looks like we have been following the Gitflow convention of repositories, where there is a master branch containing only the code releases.

Our tentative plan is to convert the VIVO Wiki on SourceForge into one or more Confluence spaces, on a system hosted by DuraSpace.
Question – does Confluence support OpenId? If so, a person could use OpenId on both Github and a Confluence wiki.

More details on the developer's list, as they become available.

Notable Development List Traffic

- Apache configuration -- alias and paths
- Failed to rebuild search index - *Brian Caruso: There was a problem parsing the JSON from the RDFService into a result set. What is odd is that /f is a valid escape in JSON for a form feed. If this is the problem it could be fixed in com.hp.hpl.jena.sparql.resultset.JSONInput.toJSON. A workaround would be to not have any form feed characters in your literals.*

- Modtime in VIVO – *The property vitro:modTime is no longer supported and I do not think that it gets recorded in the triples. There is a field in the solr index named timestamp that has the time when the individual was last updated in the solr index. (note that this Solr index property reflects changes to anything indexed for the Solr 'document' corresponding approximately to a VIVO page, so could be very useful as a way to detect when a page needs to be refreshed in an apache caching solution like Squid)*
- Special character can't be removed from Jena model – see NIHVIVO-3994
- New site, new namespace - *I wanted to change the "individual" part of the uri to just "uri" ...*
- Creating relationships in EditConfigurationGenerator - *I am attempting to add a relationship between a newly created subject and a data type property within an implementation of EditConfigurationGenerator.*

Open Discussion

- Eric Meeks want to know how to determine where to put things in the ontology. Examples are patents, clinical trials. Recommending the ontology list and the ontology call, so their questions are known to the whole group.

Call-in Information

Topic: VIVO weekly developer call

Date: Every Thursday, from Thursday, October 4, 2012 to no end date

Time: 1:00 pm, Eastern Daylight Time (New York, GMT-04:00)

Meeting Number: 645 873 290

To join the online meeting

- Go to https://cornelluniversity.webex.com/cornelluniversity/j.php?ED=161711167&UID=1278154642&PW=NMWM4ZDNhYWZm&RT=MiMxMQ%3D%3D
- If requested, enter your name and email address.
- Click "Join".

To view in other time zones or languages, please click the link: https://cornelluniversity.webex.com/cornelluniversity/j.php?ED=161711167&UID=1278154642&PW=NMWM4ZDNhYWZm&ORT=MiMxMQ%3D%3D

To join the audio conference only

- To receive a call back, provide your phone number when you join the meeting, or call the number below and enter the access code.
- Call-in toll-free number (US/Canada): 1-855-244-8681
- Call-in toll number (US/Canada): 1-650-479-3207
- Global call-in numbers: https://cornelluniversity.webex.com/cornelluniversity/globalcallin.php?serviceType=MC&ED=161711167&tollFree=1
- Toll-free dialing restrictions: http://www.webex.com/pdf/tollfree_restrictions.pdf

Access code:645 873 290