

Implementation and Development Call 20140417

Calls are held every Thursday at 1 pm eastern time – convert to your time at <http://www.thetimezoneconverter.com>

Announcements

- Ontology Working Group: [next call is Thursday, May 1 at noon EDT/ 9 am PDT](#)
 - addressing persistent VIVO URIs and related issues
 - today's call was focused on datasets and dataset ontologies, including options to engage with a W3C working group active in that area
 - The Health Care Life Sciences working group new draft standard for dataset description
 - We're trying to think how VIVO can contribute
 - Submitting use cases on how we represent datasets in VIVO
 - Patrick West: DCAT ontology is what we're using at RPI for the Deep Carbon project. Also have our own DCO ontology that we'd be happy to share. Integrating with CKAN for data flow and management.
 - Making people aware that VIVOS will have extensive contextual data around datasets
 - Also discussion of how VIVO relates to ORCID iDs and the distinctions between an ORCID iD, a VIVO URI, and the set of statements about a person in a VIVO
- Apps & Tools Working Group: [next call is April 29 at 1 pm EDT](#)
 - Find links to YouTube videos of all previous calls there
 - On April 15 call: Chris talked about FundRef and funding org data and stable URIs -- match with citizen(something).org and then visualize this
 - Alexandre Rademaker (IBM Brazil) and getting specialized data into RDF and then into VIVO
 - Apps and Tools workshop at the conference. Looking for participants to do demos -- looking for the best ways to create, use, visualize VIVO data and would love to have additional authors and help
- [VIVO Bootcamp at ELAG 2014](#) (June 10 at University of Bath, United Kingdom)
 - Violeta Ilik from Texas A&M will be there
 - the organizers (including Violeta) are looking for another representative from the VIVO community

Upcoming Activities

- First annual survey of VIVO sites
 - see Paul's email to the lists: [What does your site care about? Participate in the 2014 VIVO Annual Survey](#)
 - The survey form is live: <https://www.formstack.com/forms/?1704676-glMWIsYzom>
 - A presentation proposal has been submitted to the conference and in order to have time to compile and analyze the data, responses are requested by May 29
- Next themed weekly call topic – April 24th: VIVO End User Documentation – for supporting end users who will be editing in VIVO, whether individually or as proxy editors for a whole department (requested by University of Florida) -- can we make templated documentation where individual sites can put in tips or extra information? Rumor has it there's a good document underway -- can this be shared on GitHub using GitHub pages? (Alex: GitHub being used by non-coders for local political campaigns and other collaborative writing projects)
 - Format of themed calls could be a mix of presentations and working
 - Note that we are seeking volunteer facilitators for each themed call -- let Alex, Paul, or Jon know directly or via the listserv(s)

Updates

- Colorado (Don)
 - focused on effort to roll out Orcid iD's to our Faculty and PRA's
 - sending a pre-registration email next week
 - collecting those over the summer and will be added to the Faculty Information System (FIS) database
 - inclined to wait for 1.6 since the ontology changed slightly in its representation of ORCID iDs with 1.6
 - looking at upgrading to VIVO 1.6 prior to ingesting Orcid iD's into VIVO
 - big challenge will be to upgrade our harvest process. If there are any links to harvester code that has been converted to 1.6 I would like to see this to help bootstrap my process.
 - John Ferreira <jaf30@cornell.edu> has volunteered to be the VIVO Harvester project maintainer
 - LASP - ontology complete.
 - Other people seem to be using other tools -- have other organizations done the upgrade with other tools?
 - a mix -- the Harvester can be seen as a framework designed to have modular steps where the code used for each step may be Python just as well as VIVO
 - mention of BESSIG April 16 meeting with LASP VIVO team? Attendees from local labs including a USGS (?) group
 - LASP is thinking of putting up a SPARQL endpoint
- Cornell (Jon, Jim, Brian, et al)
 - VIVO 1.6.1 released -- not a lot of feedback yet
 - We have a number of other projects that may use all or elements of VIVO
 - Working toward a 1.7 release, keeping it modest, to get on a schedule of two releases a year in accordance with discussions at the DuraSpace Sponsor Summit in DC in early March
 - May and November seem to be the logical times to not get involved in the summer vacation and winter holiday periods
 - looking for a code freeze the end of May
 - Question: will there be major ontology changes with 1.7?
 - Answer: No. Just any bug fixes or improvements/clarifications that we receive feedback have been problems with 1.6
 - We do hope to produce the VIVO-ISF modules for 1.7 from the master ISF ontology in a more automated way, as is being worked on for eagle-i
- Duke (Patrick)
 - Question about how to add book or artistic work reviews -- helpful reply from Michaelleen (using "featured in" predicate) -- starting to look into how to represent this as "review of" (BIBO) in VIVO, as well as in Elements

- Upgrading Duke's Ruby-based "Mapper" interop to Elements 4.6 API
 - Released core of "Mapper" and Duke ontology extensions on GitHub -- will send some info to list -- contact them with questions -- started this process at Hackathon
 - https://github.com/OIT-ADS-Web/vivo_mapper
 - https://github.com/OIT-ADS-Web/duke_vivo_ontologies
- Florida (Chris)
 - Starting our 1.6 migration planning -- hoping to be migrated over before the Conference
 - Have a lot of data ingests built around the 1.5 ontology that need some re-working
 - Presented 2 posters (on figshare?) at American Medical Informatics Association (AMIA) Clinical Research Informatics Meeting.
 - dChecker (on github) - Sem Web data quality tool that is based on SPARQL and python.
 - <https://github.com/nrejack/dchecker>
 - Enterprise Data 2 RDF - PeopleSoft to UF VIVO RDF
 - have started to understand more of the test cases that have to check for in the ingest -- dual appointments, etc
 - Working on universal list of funding organizations (aka, NIH, Sloan, etc.) with URI's and geospatial location data and IRS non profit data associated.
 - Wants to create a map showing the organizations funding active grants at UF
 - About 6,000 funding organizations have DOIs, and each of them has an EIN (employee identification number) from the IRS -- an important linking variable since is in FundRef also
 - Link to Mike's Python tools on GitHub <https://github.com/mconlon17>
 - Apps & Tools group has been recruiting presenters -- not necessarily just on finished products -- starting to have working discussion sessions
- IBM Research and FGV (<http://www.fgv.br>) (Alexandre)
 - Interested in VIVO+CKAN (datasets and people, along the lines of what the Deep Carbon Observatory at RPI has done). IBM Research and City of Rio de Janeiro project can be a very interesting use case.
 - Involved in the implementation of VIVO at FGV (unfortunately, the VIVO instance is not yet public. it has 23k people and 72k research instances). Problems and questions about VIVO Performance! -- a good topic for a themed call in May
- Memorial University (Max)
 - have been trying to find documentation on migrating from 1.5 to 1.6 -- to be able to migrate their 1.5-era Knowledge Mobilization Ontology to be compatible with 1.6
 - had a session on Friday at the I-Fest with Melissa and Brian -- but Max's hard drive crashed subsequently so is looking for notes
 - Brian -- can re-send the ontology file that were working on in Protégé, and instructions about how to download the ISF again
 - Yaffle update -- have been communicating with University of New Brunswick (UNB) and CNA (College of North Atlantic) about sharing the Yaffle code
 - Thinking about separate instances for now
 - Have a front end design and have been doing usability tests
 - Have a Drupal front end that has been hooked up to VIVO, using MilesW's Drupal RDF extension (link?)
- Smithsonian (Alvin)
 - Meeting with Institutional stakeholders/funders to address their specific needs
 - Identifying data sources for import and writing RFP for contractor to import and establish method for updating/reporting per user needs
 - Jon: DuraSpace VIVO project starting to work on list of Registered Service Providers -- please let any contractors interested in VIVO work about DuraSpace's RSP directory.
- Symplectic (Alex)
 - attended UCSF-hosted Profiles developer call last Friday
 - joined SHARE Technical Working Group ([SHARE webpage](#)) -- attended call this past Tuesday
- RPI (Patrick)
 - Upgraded from 1.5 to 1.6 and migrated to a new machine all in the same step.
 - Dealt with this issue: <https://jira.duraspace.org/browse/VIVO-653>
 - Dealt with this issue: <https://jira.duraspace.org/browse/VIVO-711>
 - Applied our changes to the code (CKAN and Handle integration)
 - Changed the URI of the DCO ontology. This caused an issue with the JENA index. Basically it was corrupt. So we would have to query using the old URI <<http://deepcarbon.net/ontology/schema#projectAssociatedWith>> but the result of the query would show the new URI <<http://info.deepcarbon.net/schema#projectAssociatedWith>>
 - Basically had to export all the RDF as n-triples (ingesting RDF/XML is broken in VIVO 1.6). Dropped the mysql database, created a new empty database, re-installed VIVO, started tomcat and let it fill in. The ingested the n-triples file. It's still working (105,000 statements to work through)
 - Integration with CKAN broke with upgrade to 1.6. Fix to be installed later today
 - Still trying to work through the ontology changes in 1.6. The obo ontology is the focus of this. Since the predicate URIs are, um, not as self explanatory as you usually find.
 - Issue on Tuesday with MySQL creating massive temporary files (16GB each) given more complex SPARQL queries. This caused a file system full issue that had to be resolved by changing the MySQL configuration to point to a different temp directory.
 - Still having issues with Server Status ... huh ... actually, not seeing those issues at the moment. Perhaps it was tied to the index problem we were having.
 - Single Sign-on integration using shibboleth is completed in the test environment. We're running tests now and hope to have in production by the end of the month.
 - We now have over 1000 individuals, and each one of those will be able to log in to VIVO and add/modify content given a certain set of permissions.
 - We're starting to create some really, really cool visualizations using data from VIVO using our faceted browser, using Drupal, Google Maps and Google Earth, and using simple web pages.
 - Don in chat window: Hi Patrick - have you had a chance to sit on the VIVO ontology calls? We talked about the HCLSDataset w3c working group. Curious about your dcat extensions and this: <http://htmlpreview.github.io/?https://github.com/joejimbo/HCLSDatasetDescriptions/blob/master/Overview.html>
- UCSF (Eric is out of the office today)
- Virginia Tech (Keith)
 - Implementing Symplectic now and about 4 weeks out from getting data ingested into VIVO -- starting with 1 or 2 departments but are planning to expand to the whole of Virginia Tech
- Weill Cornell (Paul)
 - Performance
 - Still struggling with downtime and slower performance of larger profiles
 - Brian was able to get our code/data performing well on his instance

- Goal #1: test Brian's instance with our load-testing tool
- Goal #2: try installing VIVO on a dedicated server
- Goal #3: install VIVO on a dedicated VM at Cornell Ithaca for performance testing so that we have an instance instrumented for test data coming from Weill and other orgs
- Where do sites store their official organizational hierarchy including official name, type, sub/super organizations, leadership, etc.?
 - Chris -- at UF, VIVO is now the only place where there is a complete organization hierarchy
 - Believe this is the case at Brown as well -- dealing with research institutes and programmers

Notable list traffic

- [sometimes when you log in you get a bunch of status warnings -- fixed by dropping MySQL VIVO DB](#)
- [ingesting RDF already modeled with VIVO Ontology](#)
- [VIVO Usability Testing](#)
- [Very, very odd behavior with our instance data -- corrupt Jena index and had to drop the database and reload the RDF](#)
 - a complaint about the XML prefix when trying to load an RDF/XML document in 1.6 -- is this a known problem? Jim -- there was a problem in the develop branch, but should be fixed in 1.6.1
 - problem manifests that when uploading files, settings in form fields selecting formats other than RDF/XML are ignored
 - this sounds like it may be a different problem -- if you see the exception again, please send a screenshot
- [Exception during index building](#)
- [ingest RDF problem. NullPointerException](#)
- [ingest problem: things disappear after recompute](#)
- [map of science](#)
- [Update level of the obo:realizes property](#)
- [Harvester Diff Problem](#)
- [Book reviews / interviews of a scholar](#)
- [Ontology URI change](#)
- [citations](#)
- [URLs](#)

See the [vivo-dev-all archive](#) and [vivo-imp-issues archive](#) for complete email threads

Call-in Information

- Date: Every Thursday, no end date
- Time: 1:00 pm, Eastern Daylight Time (New York, GMT-04:00)
- Meeting Number: 641 825 891

To join the online meeting

- Go to <https://cornell.webex.com/cornell/e.php?AT=WMI&EventID=167096322&RT=MiM2>
- If requested, enter your name and email address.
- Click "Join".

1. Call in to the meeting:

1-855-244-8681 (Call-in toll-free number (US/Canada))

1-650-479-3207 (Call-in toll number (US/Canada))

2. Enter the access code:

641 825 891 #

3. Enter your Attendee ID:

8173 #