

Implementation and Development Call 20130919

- Calls are held every Thursday at 1 pm eastern time (GMT-4 in daylight savings, GMT-5 standard time) – convert to your time at <http://www.thetimezoneconverter.com>

View and edit this page permanently at <https://wiki.duraspace.org/x/r-AQAg>, or use the temporal [Google Doc](#) for collaborative note taking during the call.

VIVO is hiring!

DuraSpace is seeking a dynamic and entrepreneurial Project Director for the open source VIVO project (www.vivoweb.org), a world-wide community focused on creating software tools, ontologies, and services. The VIVO Project Director will have the opportunity to play a major role in a collaborative movement that will shape the future of research.

[See full posting](#)

Release update

Report on merging the dev-ISF branch into the main dev branch (in progress)

Full-scale testing should begin week after next

Apps&Tools Group

Notes from Sept. 10 meeting and WebCast: www.youtube.com/watch?v=9-95XzYs8ME

Next meeting (**timed for international attendance**): Tuesday, September 24 at 8 AM Eastern, 10 pm Melbourne, 1 pm London, 2pm Amsterdam and Rome

Agenda

1. Presentation By Nicholas Rejack on his Python data integrity checker that uses python and SPARQL in an automated fashion to keep data clean in VIVO. He calls it dChecker.
2. Updates on Working Group Goals
3. Updates on Duraspace wiki for the working group
4. Updates on curating the list of existing tools and authors/maintainers in the wiki.
5. Update on the SPARQL example templates with ontology diagrams and code samples
6. Update on Chris Barnes's work on a VIVO vagrant virtual machine that can be used by the Working Group to develop apps and tools.
7. Other business.

Updates

- Colorado (Alex) -- still making progress on Elements curation despite recent flooding issues in the area; trying to get publications information brought into their faculty reporting tool as part of the annual reporting process so that the faculty can review the data for correctness at the same time they do other updates; publications would come over to VIVO in early 2014.
- Cornell (Jon, Tim, Huda, BrianL) -- working on the release ... have an Elements instance up with their people in it, starting to explore it
- Duke (Sheri & Patrick) We continue to develop a feature to load artistic works. Have been trying to configure Elements so that data entry will take place there, and we will use our existing framework to import artistic relationships and works. We're now considering other alternatives however because Elements has the 'role' attribute on the artistic work. We need role to be on the relationship. If we are unable to add custom fields to the relationship in Elements, we will not be able to use Elements for the artistic works feature and would have to develop another system for that data entry. We're also modifying our grant import process to present a larger set - we have significantly restricted the number of grants we load into Vivo because we need to be careful not to import those that should not be publicized.
 - Can manipulate the data from Elements as bring it across, but because the role is on the work, everybody connected to a publication would be assigned the same role. Need roles beyond the current author, editor, and translator supported in Elements.
 - Would be good to bring up on the Symplectic North American user forum
 - Colorado is working with Elements but so far has been focusing on more standard publication types
- EPA (Zac) -- Cleaning up data and waiting for approval to release. Nothing new. Have a system that runs HR updates that are trying to wire up so changes get propagated to VIVO.
- Florida (Nicholas and Chris)
 - URL rewrite from <http://vivo.ufl.edu/display/cpb>
 - to <http://vivo.ufl.edu/display/n64866> using a file that maps "cpb" to n64866 and then allows apache to do the URL rewrite with a rewrite rule.
 - In `apache/sites-enabled/default-ssl`
 - RewriteMap glidmap `txt:/etc/apache2/vivo_glid_map.txt`
 - #rewrite engine changes
 - RewriteEngine On
 - RewriteMap glidmap `txt:/etc/apache2/vivo_glid_map.txt`
 - RewriteCond `/${glidmap:$1|Unknown} !Unknown`
 - RewriteRule `^/individual/(.*)$ /individual/${glidmap:$1|$1} [R,L]`
 - RewriteCond `/${glidmap:$1|Unknown} !Unknown`
 - RewriteRule `^/display/(.*)$ /individual/${glidmap:$1|$1} [R,L]`
 - "glidmap" is a file in `/etc/apache2/vivo_glid_map.txt`
 - format = `cpb n12345`
 - tab delimited file with 2 columns "label" and "vivo N number from URI"

- We have a nightly job that looks through vivo and parses out peoples email address (IE, cpb@ufl.edu) and then the N number (n12345) from the profile URI and then WRITES that to the map file in /etc/apache.
 - Mostly working on dChecker software that runs a set of SPARQL queries every day on a cron job, which not only finds malformed data but makes it easier to track down the causes of some of these data errors
 - Ran a new people ingest process that created a 52Mb RDF file of additions -- this new process works with people and contact information from Peoplesoft to decide what has changed and process the additions and retractions in VIVO; with 1.6 this can become more automated. Represents a couple months of work.
- Memorial (John, Max) -- have been wondering whether it's possible to have a standardized format for the VIVO URIs. Florida has a mapping implemented via Apache URL rewrite rule and a big matching list that allows dynamically connecting from the institution's GatorId to the native VIVO URI. Also playing with the SPARQL query builder and experimenting with SPARQL constructs. Also looking at experimenting with a Drupal front end based on what Miles Worthington has done in the past (e.g., <http://impact.cals.cornell.edu>)
- NYU (Yin) -- sympathetic to the Boulder area for its water issues ... (Alex) Thankfully the data center stayed up -- more the employees and general public affected.
- UCSF (Eric) -- OpenSocial gadgets can read but not yet produce linked open data, but are working on solving that -- extended the VIVO ontology with an ORNG space (<http://www.orng.info>). Now when gadgets save data it will be exposed as linked open data. A lot of websites are grabbing data out of UCSF Profiles and see things like news, slideshare and youtube videos, etc. and haven't been able to share that data. Each gadget would have a mini-definition of an OWL file that can formally extend the data model -- not sure how brave we want people to be in extending the ontology, but it should at least be less intimidating to play around with making data that can also be available.
 - Also, trying to use the disambiguation service built into Profiles that could be used to seed a registry of URIs for authors from PubMed -- this is the central service that could be connected to a VIVO through its API. Each week when connect to the service to find authors for UCSF, they could also indicate the local UCSF Profiles URI for authors who are identified, aligned with the Pubmed Id. This would enable other people who connect to the service to identify a URI for authors they may not be able to identify since the authors work at a different institution from their own. Titus Schleyer has connected this service to Digital Vita at the University of Pittsburgh.
 - Eric will be working with Griffin Weber to see if this can be added to the service where it's hosted at Harvard
 - Will bring this up on the tools group -- will have a very low barrier to entry. Chris would love to have Eric present the disambiguation service, even before it can be made to serve as a URI registry.
- Virginia Tech (Curtis) -- in the process of trying to set up hardware through central IT
- Weill Cornell (Paul and Eliza) -- still working on the same things as last week

Notable list traffic

- Newbie installation specs questions from Lukas at University of Amsterdam, with plans to start with a small dataset
 - Would one server with 2GB RAM be enough for this test installation? Should be enough for a test instance using a small dataset.
 - Is MySQL 5.1 required, or could MySQL 5.0.95 also work? We can't recall any changes with 5.1 that would be different for Jena's SDB library (that VIVO uses) vs. 5.0.95
 - Same for Java: 6? And Tomcat: 6 or 7? We recommend starting with Java 7 since Java 6 is no longer supported.
 - Jon shared his experience of purchasing small Linux box with 4GB/1TB drive for \$375 -- why constrain VIVO operating environment to such minimal RAM?
 - Cloud hosting is another option where you can spin up a test instance for a short time -- when you go to production at a large institution the costs can mount up. DuraSpace is interested in understanding what the VIVO community needs and how a cloud solution might work
- Customization on template -- Yu fixed after Tim responded -- that you're not seeing an error page suggests the generator class is being found and that the problem may be with the template. When you attempt to display the new custom form, do you get a blank page? Also, have you checked the vivo.all.log file and catalina.out files for error messages?

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".

To view in other time zones or languages, please click the link: <https://cornell.webex.com/cornell/globalcallin.php?serviceType=MC&ED=167096322&tollFree=1>

If those links don't work, please visit the [Cornell meeting page](#) and look for a VIVO meeting.

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://cornell.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