2019-05-14 - VIVO Development IG

Date

14 May 2019

Call-in Information

Time: 11:00 am, Eastern Daylight Time (New York, GMT-04:00)

To join the online meeting:

- Go to: https://duraspace.zoom.us/j/823948749
- Or iPhone one-tap :
  - US: +14086380968,,823948749# or +16468769923,,823948749#
- Or Telephone:
  - Dial(for higher quality, dial a number based on your current location):
  - US: +1 408 638 0968 or +1 646 876 9923 or +1 669 900 6833
  - Meeting ID: 823 948 749
- International numbers available: https://duraspace.zoom.us/zoomconference?m=Qy8de-kt6W4fMMDQCAV_3qfH1W-lxAo5

Slack

- https://vivo-project.slack.com
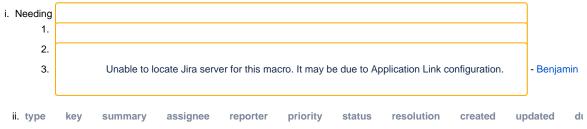 - Self-register at: http://bit.ly/vivo-slack

Attendees

blocked URL Indicating note-taker

1. Don Elsborg
2. Brian Lowe
3. William Welling
4. Benjamin Gross
5. Huda Khan ⭐
6. Mike Conlon
7. Andrew Woods
8. Ralph O'Flinn
9. Alex Viggio
10. Julia Trimmer
11. Federico Ferrario
12. Paolo Bonora
13. Hans Harlacher
14. Damaris Murry
15. Richard Outten

Agenda

1. Cineca demonstration (slides: pptx - or - google)
2. Topics for next week:
   a. Landing Design - External Search
   b. New feature: Messaging
      i. Based on the IndexingChangeListener pattern
      ii. Potentially with RDF-Patch bodies
   c. VIVO Scholars Task Force sprint update
   d. TAMU Scholars / VIVO Scholars Entities
   e. Tickets
      i. Needing
         1.
         2.
         3. Unable to locate Jira server for this macro. It may be due to Application Link configuration.   - Benjamin

      ii. | type | key | summary | assignee | reporter | priority | status | resolution | created | updated | du |
          | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Tickets

1. Status of In-Review tickets

type	key	summary	assignee	reporter	priority	status	resolution	created	updated	due

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

2. Received

type	key	summary	assignee	reporter	priority	status	resolution	created	updated	due

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

   a. Unable to locate Jira server for this macro. It may be due to Application Link configuration.

   b. Unable to locate Jira server for this macro. It may be due to Application Link configuration.

   c. Unable to locate Jira server for this macro. It may be due to Application Link configuration.

   d. Unable to locate Jira server for this macro. It may be due to Application Link configuration.

3. Bugs (1.11)

type	key	summary	assignee	reporter	priority	status	resolution	created	updated	due

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Recording

http://bit.ly/vivo-cineca-2019-05

Notes

Draft notes in Google-Doc

Cineca demonstration

Faceted browsing using Solr. Facets on the right side of the page. Clicking on entity type links (e.g. Person) on front page leads to search results with facets visible. Facets seem to include both high level class groups as well as options such as department, affiliations, roles, keywords for skills, etc.

Tabs on profile page also include ability to search the properties/values list as well as order by ascending or descending mode. Pagination is also included under the property/value list. (In VIVO speak, these are the property group tabs, and for an individual property, you can search, sort, and page through results).

Publications example shows ability to further filter publications list (and all entities).

When searching liver, why is the first result at the top? VIVO's AllText field (where all keywords from different properties are added) brings this person back but if you click on a link under the person result, the system shows the result executed across all the entities related to the person, showing that this person has many publications.

Andrew: Three features that are being demoed?

1. Faceted browsing
2. Paginated list
3. Match found

Mike: Question regarding search ordering. In discussing VIVO with other systems/profiles, discussion about ordering criteria. Has there been concern /question regarding order of search results and how that order was determined?

Cineca: The more a researcher has done, the bigger the alltext field becomes. The accuracy of the search based on alltext alone may not give us the right results. The standard approach of Solr, the last person (Carina) would have been at the top and the first result (Rossi Giorgio) would have been last. Don't use the standard approach that use the dimensions of the AllText field.

Mike: With standard approach, results may be difficult to explain or not highlight people with most scholarly works first in the list.

Julia: Not reported by faculty but did notice it in the search page. Noticed the default behavior with AllText field showed a lot of people like students first in the list.

Damaris: Did same as Cineca. Did not take into account size of AllText field but rather number of occurrences.

Ralph: Had different ways for sorting. Usually by date (default), newest to oldest. Also have a favorites flag, where those are sorted to the top automatically when looking at individuals. When looking at all people, sort faculty to the top and everyone else sorted out by order in RDF.

Faceted browsing: created a custom configuration file that defines the facets and that is used by code that extends VIVO to populate the search index with these facets.

Andrew: Other implementations of faceted browsing.

Cineca: Code is from scratch also for faceted interface. All the theme of our VIVO starts from Clarivate Responsive demo theme.

Andrew: Internationalization.

Cineca: With old interface, problem is don't have translations for all the labels. Otherwise do have support for internationalization. Don't have internationalization for the content at the moment. Not only for faceted browsing and also for fields, have properties with more than one translation in more than one language. Right now, VIVO shows both properties (both languages?) Not able to show field based on language.

Ralph: Part of internationalization work was to bring out the content that is hard-coded.

Andrew: Fedrico's example of CV available in two different languages

Ralph: Don't think covered ground for including multiple languages at an individual property level

Andrew: Making these changes available for technical discussion with an eye to making them available in the community code base. What are your thoughts with that in mind? Should we focus on one or two of these features? Difficult to add lots of code at once and easier to review smaller pieces of code. Possible to break up these features and introduce the updates in smaller incremental ways so it's easier to put into code base versus a massive pull request that includes everything.

Cineca: Possible. Better also for us to work on one feature at a time. Simplest is probably paginated list. Most interesting probably is faceted browsing or match found. Can decide together.

Andrew: People see barriers why code would not be generalizable or generally applicable.

Ralph: If code accessible, would rather look at it and try to understand it.

Cineca: On version 1.9.2. We use blazegraph and some problem with this kind of back-end. With pull request, we fixed this problem and can move to a newer version. Use BlazeGraph with Jena but not MySQL (instead RDF native database). Load with Pentaho because don't use editing part of VIVO because have CRIS system where publications (DSpace), and others (resource module). CRIS for loading data and then push into VIVO using our ETL based on Pentaho Data Integration (also known as Kettle) .

Huda: Interface features for faceted browsing and search through properties are great. Search relevance is a more involved issue so as long we as keep that configurable/testable would be good (so each installation can define what relevance means to them and how to test it). With respect to code, would need to review but if it is extending the code within VIVO itself, seems like a good candidate to review.

Version 1.9.2

They use blazegraph so this is limited to 1.9.2 because of some sort of bug

They use Jena with Blazegraph with VIVO?

Don't use editing part of VIVO!!

CRIS system loads data,

ETL is Pentaho/Kettle

Blazegraph - solved performance problems for loading

They load to the VIVO update API.

With Pentaho, they only load updated information, they had out of memory issues with tomcat when they did all or nothing loads

Actions

- Unable to locate Jira server for this macro. It may be due to Application Link configuration.   - Mike Conlon, can you give this one a

Previous Actions

-