

# 2018-08-21 - VIVO Development IG

## Date

21 Aug 2018

## Call-in Information

Time: 11:00 am, Eastern 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/join?m=Qy8de-kt6W4fMMDQCAV\\_3qfH1W-lxAo5](https://duraspace.zoom.us/join?m=Qy8de-kt6W4fMMDQCAV_3qfH1W-lxAo5)

## Slack

- <https://vivo-project.slack.com>
  - Self-register at: <https://goo.gl/forms/JxQFkut4TYj4Ehww1>

## Development Process

- See <https://github.com/vivo-project/VIVO/wiki/Development-Processes>

## Attendees

[blocked URL](#) Indicating note-taker

1. [Ralph O'Flinn](#)
2. [Jim Blake](#)
3. [Huda Khan](#)
4. [Tim Worrall](#)
5. [Andrew Woods](#)
6. [Mike Conlon](#) ★
7. [Christian Hauschke](#)
8. [Alex Viggio](#)
9. [Brian Lowe](#)
10. [Benjamin Gross](#)

## Agenda

1. Report from the field: [ElasticSearch instead of Solr](#) (Jim Blake )
2. [Documenting ingest approaches](#)
3. Sept [sprint planning](#)
4. Active tickets
  - a. Unable to locate Jira server for this macro. It may be due to Application Link configuration. (pending response - [Benjami](#)
  - b. Unable to locate Jira server for this macro. It may be due to Application Link configuration. ([Muhammad Javed](#) - to
  - c. Unable to locate Jira server for this macro. It may be due to Application Link configuration. ([Kitio Fofack](#) - where does
5. [How to and walk-through of:](#) Unable to locate Jira server for this macro. It may be due to Application Link configuration.
6. [Search index](#)
  - a. [Search index](#)
  - b. [Trinlosters](#)
7. Unable to locate Jira server for this macro. It may be due to Application Link configuration.

# Notes

Draft notes in Google-Doc

## ElasticSearch instead of Solr

1. Swapping in ElasticSearch 6.3 for Solr.
2. Jlm demonstrated VIVO 1.8 running elasticSearch.
3. Downloaded elasticSearch last week and spent a couple of days creating a VIVO driver.
  - a. See [https://github.com/j2blake/Vitro/blob/feature/elasticsearchExperiments/api/src/main/java/edu/cornell/mannlib/vitro/webapp/searchengine/elasticsearch/Elasticsearch\\_notes\\_on\\_the\\_first\\_draft.md](https://github.com/j2blake/Vitro/blob/feature/elasticsearchExperiments/api/src/main/java/edu/cornell/mannlib/vitro/webapp/searchengine/elasticsearch/Elasticsearch_notes_on_the_first_draft.md)
  - b. Looks like VIVO. Works like VIVO.
4. Needs:
  - a. Docs
  - b. Smoke test
  - c. Functional testing
  - d. Improved snippets
  - e. Code improvements
  - f. Unit tests
  - g. Automatic initialization of the index
5. Why Elasticsearch
  - a. Create options for sys admins.
  - b. Some sites are already bought in to ElasticSearch. They love it.
  - c. Certainly should not run in parallel. Should be implemented as part of VIVO.
6. Put more data in the index
7. Elastic has nested fields that keep their relationships publication uris on the author's record. Publication names.
  - a. Does the current version of Solr have these features?
8. Ownership characteristics – both seem okay? VIVO rather small by elasticSearch standards?
9. ElasticSearch has an Apache license
10. Both ElasticSearch and Solr are based on underlying Lucene technology
  - a. See In case this helps (not sure if the article is 'good' or not but here you go): <https://www.searchtechnologies.com/blog/solr-vs-elasticsearch-top-open-source-search>
  - b. And <https://sematext.com/blog/solr-vs-elasticsearch-differences/> and <https://db-engines.com/en/ranking/search+engine> (rank Elasticsearch #1 based on their ranking methodology)
11. Perhaps not a search for "best" but which is best in a particular environment and for a particular application
12. After Solr 5, (current is Solr 7) Solr is a free-standing application, like ElasticSearch.
13. Explore nested documents in elasticSearch. Current VIVO interface
14. Developed in 1.8 since the ant environment is so much more productive than the 1.9/1.10 Maven environment.
15. Are there touch points with product evolution?
  - a. Does the nested doc capability, and the ability to have ElasticSearch in the architecture resonate with product evolution?
  - b. Product Evolution is looking at GraphQL for its API capabilities.
16. How do we see applications related to VIVO being installed? By the installer?
  - a. For try out, a jar?
  - b. For try out, a VM?
  - c. For production, apps must be installed? As we require MySQL and Tomcat now? Solr or ElasticSearch in the future?

## Ingest Approaches

1. Add ingest tools to the table in the apps and tools catalog

## Topics for September Sprint

1. abox/tbox topic
2. ElasticSearch
3. Internationalization
4. Decoupling?

## Planning for walk-through of large pull-request

1. Graham's pull request is a big one. May need some additional hands.

## Previous Actions

- [Don Elsborg](#) to document "firsttime" resolution in CU Boulder wiki, circulate this doc to email list and discuss as a team how to integrate this in VIVO documentation
  - moved initialTBoxAnnotations.n3 back to firsttime. Need to edit this to change a few labels, eg authors "CU Boulder Authors"
    - eg: <https://github.com/cu-boulder/vivo-cuboulder/blob/a3b7e70704264b2c4fead29ef80c5efc45a9357e/home/src/main/resources/rdf/tbox/firsttime/initialTBoxAnnotations.n3>
  - Had some changes in propertygroups.rdf, in 1.7 this was moved to firsttime. Left it there now

## Actions

- [Alex Viggio](#) will bring news of Elasticsearch instead of Solr up with Product Evolution. Might there be consequences for the September sprint.

