

2018-09 Sprint

Dates

Sept 17-28, 2018

Participants

1. [Graham Triggs](#)
2. [Andrew Woods](#)
3. [Mike Conlon](#) (Second week: September 24-28, documentation, testing)
4. [Qazi Asim Ijaz Ahmad](#) (Not sure about the dates yet. Would like to work on Elastic Search with Don)
5. [Kitio Fofack](#) mainly on i18n (available for the sprint)
6. [Christian Hauschke](#) (2nd week, i18n)
7. [Ralph O'Flinn](#)
8. [Brian Lowe](#)
9. [Don Elsborg](#)
10. [Steven McCauley](#)
11. Manuel Schwarz

Meetings

[2018-09-17 Sprint - Kickoff Meeting](#)

Tickets

[Scrum Board](#)

Elasticsearch / Solr Upgrade

Team

1. [Don Elsborg](#) (CU Boulder)
2. [Qazi Asim Ijaz Ahmad](#)
3. [Steven McCauley](#)
4. [Ralph O'Flinn](#)

Purpose/Rationale

1. Security concerns with Solr dependencies
2. Integrated elastic index in VIVO to support facetview UI used by Unavco, CU Boulder, DCO
3. Brings index work from Product Evolution group closer to the VIVO core
4. Deliver more structured and rich json and eventually json-ld documents from VIVO's indexes for web consumers

Deliverables

1. ~~Discuss/Analyze/Document the ElasticSearch work done by individual sites.~~
 - a. ~~Which sites?~~
2. Create a design for externalized search
 - a. Support [ElasticSearch](#) (v6.4.0)
 - b. Support [Solr](#) (7.4.0 & 4.10?)
3. Implement externalized search
 - a. [Initial Solr work from Huda](#) ([Ralph O'Flinn](#), [Steven McCauley](#))
 - b. [Initial ElasticSearch work from Jim](#) ([Don Elsborg](#))
 - c. merge pull requests from items a and b above into a new combined sprint branch
 - d. verify the search engine abstraction layer such that VIVO can work with both ES and SOLR based on above 2 items
4. Create instructions on how to make YOUR VIVO installation work with ES or SOLR (on the assumption that Solr will still be the default).

Second priority deliverables

1. Analyze how to build a nested json-doc that represents an object (person, publication, grant, etc) in the index (both SOLR or Elastic)
2. Lay groundwork for analysis of incorporating semantics (json-ld, other) in the indexed document. Mapping the objects from VIVO-ISF to an indexed semantic document would need to involve the ontology group
3. Request: keeping the delivered Elasticsearch integration backwards compatible with VIVO 1.9.3 and 1.8.x – could be valuable

Relevant docs

1. <https://www.searchtechnologies.com/blog/solr-vs-elasticsearch-top-open-source-search>
2. <https://sematext.com/blog/solr-vs-elasticsearch-differences/>
3. <https://db-engines.com/en/ranking/search+engine>

Multi-Language Support

Team

1. [Christian Hauschke](#)
2. [Kitio Fofack](#)
3. [Ralph O'Flinn](#)
4. Manuel Schwarz

Purpose

Implement interface i18n on home page, the capability Map and the profile header form.
The interface should be able to completely switch to the selected language.

Deliverables

1. ~~Identify all places in code/freemarker/jsp that need extraction and create related JIRA Tickets~~
 - ~~Split codebase across different people (per package?)~~
 - ~~ACTION: Kitio to divide codebase~~
2. Initialize language artifacts necessary to be complete for interface i18n
 - a. Create the translations (English, German, French)
3. Improve toggling mechanism (do not use flags)
4. Improve technical documentation for adding new languages
 - Pulling in Maven artifacts

Future deliverables

1. Rectifying differences in grammar
2. Pluralization - need appropriate infrastructure
3. i18n support for multi-language content

ABox / TBox RDF Loading

Team

1. [Don Elsborg](#)
2. [Brian Lowe](#)
3. [Mike Conlon](#)

Purpose

1. Identify and understand current documentation
 - a. [Accessing VIVO Data Models#InitializingtheModels](#)
 - b. [Directories and Files](#)
 - c. [Graph Reference](#)
 - d. [Ontology Reference](#)
2. Understand and Document recommendations for updating and overriding `firsttime` and `everytime` RDF files
3. Understand and Document how this RDF data is organized in the content and configuration triple stores

Deliverables

1. Documentation of Content, Structure, Purpose, etc of 'firsttime' / 'everytime' / 'filegraph' abox,tbox RDF files and each of the directories that they're located in. See item 1.b 'Directories and Files' above
2. Documentation of where 'firsttime' / 'everytime' / 'filegraph' abox/tbox RDF is loaded (which triplestores-config vs content stores), and why
3. Verify the documentation of the relationships with vdata/rdf and the actual graphs – changes in the GUI don't get reflected in \$vivo-home/rdf
4. Documentation of pitfalls – things not to delete. (ClassGroups? And other things that are loaded by firsttime and are difficult to reset to their initial state without clearing the triple store)
5. Documentation of use cases for updating 'firsttime' / 'everytime' RDF files
6. Documentation of recommendations for above use cases

Future deliverables

1. Resetting 'firsttime' - [Brian Lowe](#) to elaborate
2. Persisting updates from the GUI (ClassGroup / PropertyGroup / etc) - [Don Elsborg](#) to elaborate