Search Index Configuration Task Force Charter

The Search Index Configuration Task Force is inactive, because Jim Blake was unwilling to continue as facilitator. The task force may be re-activated if someone else is willing to facilitate.

Context

Release 1.8 of VIVO includes a search indexer that has been substantially re-written. Compared to earlier versions, it is more efficient, easier to understand, easier to configure, and easier to evaluate.

This increased level of visibility has made it obvious that the current configuration contains errors, redundancies, and omissions. It also brings this question to the forefront: what should the search index contain? How can this best be specified?

The restructured search indexer also lays the groundwork for added functionality. In particular, we would like to investigate the work done on faceted search results at the VIVO Hackathon 2014.

Objectives

- To determine what data should be included in the standard distribution of the search index for Vitro. This should recognize that Vitro is ontology-neutral.
- To determine what data should be included in the standard distribution of the search index for VIVO. This should be based on usage patterns and the structure of the VIVO ontology. It should recognize the tradeoffs between power and performance.
- To develop a style for those configurations that is concise, easy to understand, easy to extend, and resistant to errors.
- To make code enhancements that will support that style. This part of the effort should be minor.
- To document the configuration process so individual VIVO sites can easily understand and customize VIVO search functionality to suit their needs.
- To document the changes in content from previous releases.
- To lay the ground work for configurable facets. To assess some use cases, potential roles for extensions like Blacklight, a preliminary design for the configuration file for the facets, affected classes, and jira tickets.

Deliverables:

- A configuration for the search indexer, for use in Vitro.
- A configuration for the search indexer, for use in VIVO.
- A document or Wiki page that describes to VIVO maintainers how these configurations differ from those of previous releases.
- One or more Wiki pages that describe the configuration process.
- A requirements document for configurable faceted search, including use cases, an example configuration file, survey of affected code classes, and JIRA tickets.
- A draft charter for a task force devoted to implementing faceted search.
- JIRA ticket(s) for tasks that are raised by the task force, but are considered to be beyond the scope of the current effort.

Suggested schedule:

Develop charter	1.5 weeks	2015-02-06
Solicit members and schedule first meeting	1 week	2015-02-13
Develop action items, assign to team members.	2 weeks	2015-02-27
Accomplish deliverables, preliminary review	3 weeks	2015-03-20
Revise deliverables, final review	2 weeks	2015-04-03

Members

Jim Blake (Cornell University) - task force lead

Don Elsborg (University of Colorado, Boulder) - member

John Fereira (Cornell University) - member

Huda Khan (Cornell University) - member

Jing Wang (Johns Hopkins University) - member

Ted Lawless (Brown) - member

Stephan Zednik (RPI) - interested party

Meeting Times

February 27, 2015 - 11:00 a.m. Eastern

March 11, 2015 - 2:00 p.m. Eastern

Other meetings, TBD

Communication Channels

- Distribute proposed charter and solicit participation from community via dev-all mailing list
- Announce task force creation and progress on the weekly developer/implementation calls.
- Create a page in the VIVO wiki (wiki.duraspace.org), with child pages for agenda, notes, and additional information
- Utilize JIRA to assign, document, and track specific deliverables.

Agendas and Notes

- Draft agenda for first meeting is here: Search Indexer Task Force - First meeting
  - Temporary google doc is here: http://bit.ly/1BHboGM

- Draft agenda for second meeting is here: Search indexer task force - Second meeting
  - Temporary google doc is here: http://bit.ly/1b0iD0L