

Infrastructure - Task #3142

Task # 3635 (New): Phase two Morpho implementation

Need a new local search mechanism

2012-08-17 00:00 - Jing Tao

Status:	New	Start date:	2012-08-16
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:	2013.10-Block.2.1	Story Points:	
Milestone:	None		
Product Version:	*		

Description

Currently, morpho puts all eml documents into a collection of DOM tree, then uses XPathAPI.selectNodeList method and xpath to search the DOM in the local search. If a user has a big set of local eml documents. It will cause the memory issue.

We might use Lucene search engine to replace the current approach.

Dataone uses SOLR for search. However, SOLR is a webapps. It involves tomcat, jetty and et al web servers. It is not good for morpho as a desktop application. But Lucene, the search engine used by SOLR, can be embedded into java application. So we can use Lucene.

Lucene has two main components - indexing and search.

The indexing part we might borrow some code from Dataone (mercury). I talked with Skye. We can't use it directly and have to do some modification.

The search part in Lucene is simple. It has a QueryParser class to help user to build a query from a string. Then user can call IndexSearcher.search(query) to get result.

History

#1 - 2012-10-11 15:33 - Dave Vieglais

- Target version changed from Sprint-2012.37-Block.5.3 to Sprint-2012.41-Block.6.1

#2 - 2012-10-24 18:20 - Ben Leinfelder

- Target version changed from Sprint-2012.41-Block.6.1 to Sprint-2012.44-Block.6.2

#3 - 2012-12-12 16:51 - Chris Jones

- Target version changed from Sprint-2012.44-Block.6.2 to Sprint-2012.50-Block.6.4

#4 - 2013-03-01 18:33 - Ben Leinfelder

- Target version changed from Sprint-2012.50-Block.6.4 to 2013.10-Block.2.1

#5 - 2013-03-02 05:15 - Ben Leinfelder

- Parent task changed from #3075 to #3635