Infrastructure - Bug #2617

Search by identifier in one-mercury not working

2012-04-19 17:05 - Skye Roseboom

Status:	Closed	**Start date:**	2012-04-19
Priority:	Normal	**Due date:**	
Assignee:	Skye Roseboom	**% Done:**	100%
Category:	d1_indexer	**Estimated time:**	0.00 hour
Target version:	Sprint-2012.17-Block.3.1		
Milestone:	CCI-1.0.0	**Story Points:**	
Product Version:	*		

Description

Searching by document identifier does not consistently work in one-mercury's default 'full text' search.

This problem arose due to removing the copyField directives from the the search solr schema related to the full text field. This led to the pid value no longer being present in the full text search field, breaking pid searching. Since the dataONE pid value is not present in the science metadata document, the full text field is cannot solely be generated from the science metadata document.

History

#1 - 2012-04-23 16:32 - Skye Roseboom

- Status changed from New to In Progress

#2 - 2012-04-23 17:52 - Dave Vieglais

- Target version changed from Sprint-2012.15-Block.2.4 to Sprint-2012.17-Block.3.1

- Position set to 1

- Position changed from 1 to 427

#3 - 2012-04-24 19:46 - Skye Roseboom

- Status changed from In Progress to Closed

Updated search index solr schema to copy 'id' field into the 'text' field which is used to support the fullText search feature in one-mercury. Deployed and tested on cn-dev-2.dataone.org.

#4 - 2012-04-27 19:24 - Skye Roseboom

Initial approach to use solr's copyField directive in schema.xml was not a good solution. It triggered data files to begin showing up in 'full text' search results. This was because non-science metadata documents now had text (pid) in the full text field.

Changed solution so that the solr field generation logic responsible for the 'full text' field the cn_index_processor project now adds the identifier string to the full text field value. This allows only science metadata documents to have the pid value appended to the full text field.