

Infrastructure - Task #8775

Make taxonomic rank fields in Solr index non-case-sensitive

2019-03-11 22:39 - Bryce Mecum

Status:	In Progress	Start date:	2019-03-11
Priority:	Normal	Due date:	2019-03-12
Assignee:	Bryce Mecum	% Done:	30%
Category:		Estimated time:	0.00 hour
Target version:		Story Points:	
Milestone:	None		
Product Version:	*		
Description			
In the current system, EML documents with taxonomic coverage get indexed into fields such as species if they contain XML such as:			
<pre>...snip <taxonomicCoverage> <taxonomicClassification> <taxonRankName>Species</taxonRankName> <taxonRankValue>Some species</taxonRankValue> ...snip</pre>			
The field values are extracted using the XPath in In https://repository.dataone.org/software/cicore/trunk/cn/d1_cn_index_processor/src/main/resources/application-context-eml-base.xml :			
<pre>//taxonomicClassification/taxonRankValue[../taxonRankName="Species"]/text()</pre>			
We ran into a case where the taxonRankName had been entered as 'species' instead of 'Species' and we decided that the XPath is too restrictive and that the strictness is needless and surprising. This change should result in a slight but negligible decrease in performance.			
<ul style="list-style-type: none">• Change all EML taxonomy fields to also match the lowercase form of each taxonomic rank• Check over other indexing field definitions related to taxonomy to make sure the above change is consistent			

Associated revisions

Revision 19516 - 2019-03-11 23:33 - Bryce Mecum

Make taxonomic rank fields in Solr index non-case-sensitive (refs #8775)

Revision 19516 - 2019-03-11 23:33 - Bryce Mecum

Make taxonomic rank fields in Solr index non-case-sensitive (refs #8775)

History

#1 - 2019-03-11 23:35 - Bryce Mecum

- % Done changed from 0 to 30
- Status changed from New to In Progress

Dropped this fix, with a backing regression test in [r19516](#).

I didn't mark this as done yet because I wasn't sure if this needs to get targeted for an upcoming CCI release.