

Metacat - Bug #34

text matches in query algorithm don't span TEXT nodes

06/27/2000 08:23 PM - Matt Jones

Status:	Resolved	Start date:	06/27/2000
Priority:	Low	Due date:	
Assignee:	Matt Jones	% Done:	0%
Category:	metacat	Estimated time:	0.00 hour
Target version:	Sep2000	Spent time:	0.00 hour
Bugzilla-Id:	34		

Description

When text sequences from the document exceed the max field length of the database field (4000 characters for Oracle), we break the text into multiple, adjacent TEXT nodes of 4000 characters each, with the last node containing the remainder. Our current search algorithms only will find phrases within a single TEXT nodes, so if the search term spans the node boundary, the search algorithm will not return a hit. I expect this is extremely rare, basically inconsequential, because text nodes > 4000 characters will be rare, and then the search phrase has to cross the node boundary, both of which are improbable.

History

#1 - 06/27/2000 08:26 PM - Matt Jones

Because it is improbable, will require major query algorithm changes to fix it, and will have minor impact when encountered, I am not planning on fixing this bug unless it becomes evident it is more significant than my initial assessment.

#2 - 03/27/2013 02:13 PM - Redmine Admin

Original Bugzilla ID was 34