

## Semtools - Bug #5366

### Integrate semantic query with keyword and spatial query

04/04/2011 01:53 PM - ben leinfelder

<b>Status:</b>	New	<b>Start date:</b>	04/04/2011
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	ben leinfelder	<b>% Done:</b>	0%
<b>Category:</b>	Metacat extensions	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	metacat-plugin-0.9.0	<b>Spent time:</b>	0.00 hour
<b>Bugzilla-Id:</b>	5366		
<b>Description</b>			
Combining semantic, keyword and spatial criteria will allow a smoother transition from non-semantic metadata to semantically annotated data. It will also allow us to refine query results further by limiting to specific organizations, locations, people, etc.			
Metacat should support combining "squery", "semquery" and "spatialquery" actions in an extensible way that lets future query mechanisms contribute conditions to the query so that additional (or fewer) matches are returned.			
This may require additional or wrapper syntax for specifying a compound, multi-faceted query.			

#### History

##### #1 - 04/11/2011 05:55 PM - ben leinfelder

I've changed the DBQuery class in Metacat to allow combinations of a docid list (semantic and spatial queries generate these) with normal pathquery/keyword search criteria.

You can choose the normal mode (as it was before I made this change), or UNION mode or INTERSECT mode. The last two either expand or refine the search results.

The web UI still needs some work to incorporate these new features in an intuitive way.

##### #2 - 04/13/2011 10:16 AM - ben leinfelder

There a couple different behaviors possible for compound query techniques, and I'm not sure which we should choose or whether it should be user choice.

Expanding paradigm:

1. semantic concepts WITH keywords  
-all docs matching the semantic concepts plus all docs that match the keywords
2. ONLY semantic concepts  
-all docs matching the semantic concepts plus any other non-annotated docs (keyword contains "%")
3. ONLY keyword  
-all docs matching the keywords (might still include annotated docs)

Refining paradigm:

1. semantic concepts WITH keyword  
-docs matching the semantic concepts if and only if they match the keywords as well
2. ONLY semantic concepts  
-only docs matching the semantic concepts (non-annotated documents are not returned)
3. ONLY keywords  
-all docs matching the keywords (might still include annotated docs)

##### #3 - 04/13/2011 02:08 PM - ben leinfelder

from Shawn:

I vote for expanding, but with ranking.

This is like a google keyword search, where a match simply contains one of the search words, but docs with all search words are ranked higher (obviously the google algo is more complicated than this...but this is the general idea)

##### #4 - 03/27/2013 02:30 PM - Redmine Admin

Original Bugzilla ID was 5366

##### #5 - 11/26/2013 08:56 AM - ben leinfelder

- Target version changed from Unspecified to sms-index-1.0.0

**#6 - 11/26/2013 09:10 AM - ben leinfelder**

- *Target version changed from sms-index-1.0.0 to metacat-plugin-0.9.0*