

EDIT - bug #5868

re-establish support for wildcard at start of query string (includes the query string '*')

06/06/2016 10:05 AM - Andreas Kohlbecker

Status:	Feedback	Start date:	
Priority:	Highest	Due date:	
Assignee:	Andreas Müller	% Done:	50%
Category:	cdmlib	Estimated time:	0:00 hour
Target version:	Release 5.3	Found in Version:	
Severity:	major		
Description			
Norbert wrote:			
Eine Frage noch zur * wildcard: warum kann ich jetzt nicht mehr bei der erweiterten Taxon-Search mit nur "*" suchen, aber weiterhin bei der einfachen Taxon-Search, wo es gar keinen Sinn macht?			
Bei der erweiterten Taxon-Suche erlaubte früher die Kombination aus "*" + Area, das schnelle Erstellen einer Checkliste; Wegfall dieser Möglichkeit ist deshalb ein echter Verlust!			
This regression probably is a regression due to the search framework upgrade.			
TODO:			
<ul style="list-style-type: none">• find the cause for this regression and re-enable this feature• implement tests for the various search modes. Wildcards, phrase, ...			
related to #3807			
Related issues:			
Related to EDIT - feature request #7517: Make search parameters like wildcard... New			

Associated revisions

Revision d6bdbc3d - 09/14/2018 09:30 PM - Andreas Kohlbecker

fix #5868 restoring freetext search with leading wildcard and with empty query string, also fixes bug introduced while solving #4223

Revision a0557101 - 09/14/2018 09:33 PM - Andreas Kohlbecker

fix #5868 avoiding unnecessary "complex phrase" queries, fixes bug introduced while solving #4223

History

#1 - 06/06/2016 02:23 PM - Andreas Müller

- Severity changed from normal to major

#2 - 06/17/2016 11:33 AM - Andreas Kohlbecker

- Subject changed from re-establish support for wildcard at start of query string to re-establish support for wildcard at start of query string (includes the query string '*')

#3 - 02/15/2017 01:13 PM - Katja Luther

- Description updated

- Priority changed from New to Highest

#4 - 06/28/2018 02:42 PM - Andreas Kohlbecker

- Private changed from Yes to No

#5 - 06/28/2018 02:42 PM - Andreas Kohlbecker

- Related to feature request #7517: Make search parameters like wildcards configurable added

#6 - 06/28/2018 02:48 PM - Andreas Kohlbecker

- Target version changed from Unassigned CDM tickets to Release 5.2

#7 - 08/17/2018 03:54 PM - Andreas Kohlbecker

- Target version changed from Release 5.2 to Release 5.3

#8 - 09/06/2018 11:17 AM - Andreas Kohlbecker

- Target version changed from Release 5.3 to Release 5.4

#9 - 09/14/2018 09:30 PM - Andreas Kohlbecker

- Status changed from New to Resolved

- % Done changed from 0 to 50

Applied in changeset [cdmlib|d6bdbc3d4a6d81dd3298913e1a31b64d76b88211](#).

#10 - 09/14/2018 09:31 PM - Andreas Kohlbecker

- Target version changed from Release 5.4 to Release 5.3

please review

#11 - 09/15/2018 09:39 AM - Andreas Müller

- Assignee changed from Andreas Kohlbecker to Andreas Müller

#12 - 09/16/2018 02:25 PM - Andreas Müller

- File picture426-1.png added

- Status changed from Resolved to Feedback

- Assignee changed from Andreas Müller to Andreas Kohlbecker

I get a warning when searching for Achil* on test.cyprus. This is probably not related to this ticket but mention this here as I do the review. It looks like a "sec" is required, but especially misapplications often do not have a sec.

Search results



For easier theme development, the theme registry is being rebuilt on every page request. It is *extremely* important to [turn off this feature](#) on production websites.



Notice: Undefined property: stdClass::\$sec in *compose_list_of_taxa()* (line 510 of */var/www/drupal-7.32/sites/_jenkins-workspace/modules/cdm_dataportal/includes/taxon.inc*).

[Modify search](#)

Display image thumbnails

Achillea

? *Achillea santolina* Test sensu Baumel, A.

? *Achillea santolina* Test sensu J. Chrtek & B. Slavík

Achillea santolina aaa sensu J. Chrtek & B. Slavík

#13 - 09/16/2018 02:32 PM - Andreas Müller

Wildcards in the same word seem to work. I tested with "chi*ll" on test.cyprus.

However, "chi*ll*san" to find *Achillea santolinoides* Lag. and similar taxa did not work. It needs a whitespace like "chi*ll san*". Is this expected behavior, then we should mention it in the search explanation.

#14 - 09/16/2018 02:37 PM - Andreas Müller

For names with multiple whitespaces it is even more tricky.

wilhelm returns "Achillea santolinoides subsp. wilhelmsii (K. Koch) Greuter"

However, "*subsp. *wilhelm*" does not return anything. Same with "Achil* * * *wilhelm*" or "Achil* *wilhelm*"

#15 - 09/16/2018 02:45 PM - Andreas Müller

The search seems not to be sorted. This is what I get for "V"

Vitex
Vitis vinifera L.
Vitis
Vulpia brevis Boiss. & Kotschy
Vulpia
Vulpia ciliata Link
Vulpia fasciculata (Forssk.) Fritsch
Vulpia muralis (Kunth) Nees
Vulpia myuros (L.) C. C. Gmel.
Verbenaceae
Hordeum vulgare L. subsp. vulgare
Viola heldreichiana sensu R. D. Meikle
Vulpia membranacea sensu R. D. Meikle
Aegilops biuncialis Vis. subsp. biuncialis
Lobularia libyca (Viv.) Meisn.
Verbascum geminiflorum Hochst.
Aloe vera (L.) Burm. f.
Moricondia spinosa subsp. maroccana sensu R. Vogt & A. Aparicio
Amaranthus graecizans subsp. silvestris (Vill.) Brenan
Vaccaria hispanica (Mill.) Rauschert
...

#16 - 09/16/2018 02:45 PM - Andreas Müller

was it like this before. Is the sorting by some kind of percentage?

#17 - 09/16/2018 02:52 PM - Andreas Müller

The area search has a problem. Searching for "V" and division1 returns NO result.
Searching for "".

Searching for "" and "Cyprus" return 1 1/2 pages (but should return almost the whole list)

#18 - 09/16/2018 07:09 PM - Andreas Müller

OK, on the new version of the cyprus database it works better. Especially the geo-search works better.

But synonym search does not seem to work.

Searching on "Phyla can*" with division6 and accepted+synonym finds Phyla canescens, but same search on "Phyla fil*" does not find its synonym Phyla filiformis (Schrad.) Meikle, but should.

#19 - 09/16/2018 08:43 PM - Andreas Müller

Andreas Müller wrote:

OK, on the new version of the cyprus database it works better. Especially the geo-search works better.

But synonym search does not seem to work.

Searching on "Phyla can*" with division6 and accepted+synonym finds Phyla canescens, but same search on "Phyla fil*" does not find its synonym Phyla filiformis (Schrad.) Meikle, but should.

This seems to happen only with subtree filter switched on. But is maybe related to the fix of this ticket as it was not recognized before. Looks like the current implementation does not return any synonyms if the subtree filter is switched on.

The subtree filter works very similar to the classification filter but uses a termQuery instead of IdQuery:

```
if(classification != null){
    finalQueryBuilder.add(taxonBaseQueryFactory.newEntityIdQuery("taxonNodes.classification.id", classification), Occur.MUST);
}
if(subtree != null){
    finalQueryBuilder.add(taxonBaseQueryFactory.newTermQuery("taxonNodes.treeIndex", subtree.treeIndex
```

```
Wc(), true), Occur.MUST);  
}
```

The implementation of the term query has been changed within this ticket.

#20 - 09/17/2018 03:22 AM - Andreas Müller

The classification filter does NOT have any effect in the webservice (1). But it has in the service(2)!

I tested (1) by changing the uuid in the service => returning same result, but should be empty because the classification does not exist. Also I added a second classification with a taxon name that is in both classification => both taxa were returned, though only 1 is in the filter classificaiton. Looks like the classification is set to null somewhere.

Not testing the return valud of `classificationService.find(uuid)` is definetely a bug. It should return an error message if classification does not exist. Or at least an empty result.

Then, the reason why it is always null is that the former parameter tree has been renamed to `classification(Node)Uuid` ([cdmlib|bfda0eba8f7](#) and [cdmlib|3b2b43193ee](#)). Therefore it is not effective anymore :-)

I tested (2) by writing a test in `TaxonServiceSearchTest`. Not yet committed.

With the result of (2) - having an effect - the expected behavior is, that synonyms are not found at all, because the classification filter same as the subtree filter queries "taxonNodes.xxx", but for synonyms there is no such field. The field should be "accTaxon.taxonNodes.xxx" but this is not indexed. Maybe this worke some time ago when synonyms were still `synonymRelations` and the result was handled with the `taxonRelationship` index. Currently it can't work.

#21 - 09/17/2018 04:42 AM - Andreas Müller

I moved synonym search and classification/subtree filters to new ticket [#7764](#) as it is not really related to the topic of **this** ticket

#22 - 09/17/2018 10:09 PM - Andreas Müller

- Assignee changed from *Andreas Kohlbecker* to *Andreas Müller*

I will try to write tests for the remaining issues and move them to a new ticket.

Files

picture426-1.png	18.2 KB	09/16/2018	Andreas Müller
------------------	---------	------------	----------------