

EDIT - bug #6151

implement pager for search service

10/21/2016 01:29 PM - Andreas Kohlbecker

Status:	Closed	Start date:	
Priority:	Highest	Due date:	
Assignee:	Andreas Kohlbecker	% Done:	100%
Category:	utis	Estimated time:	0:00 hour
Target version:	UTIS 1.3	Found in Version:	
Severity:	normal		
Description			
UTIS is not returning all possible results from the connected providers.			
A pager for search service should be implemented. Pagers only make sense for <i>like</i> and <i>fuzzy</i> searches.			
For UTIS being a federated search engine a general paging can be implemented in two ways:			
<ol style="list-style-type: none">1. The PAGESIZE limit defines the number of results in the whole response aggregated from the individual responses from multiple checklists. Since the number of hits per checklist provide in unknown it would be required to always asked to return PAGESIZE elements for a request. With the current architecture of UTIS this can impose a big overhead on the overall request processing which increases with the amount of clients that are queried in parallel.2. The PAGESIZE limit only affects the individual checklist queries. The unified response of UTIS contains all responses from the checklist providers. When UTIS requests n checklists the UTIS response can in this case have a maximum size of n * PAGESIZE response items.			
Strategy 2. seems to be the better option, since it is not causing performacne problems.			
Checklist provider services and paging:			
<ul style="list-style-type: none">• ENUNIS: no limitation: DONE (pager implemented)• COL (bgbm-cdm-serever[col]): via name Catalogue service seems to have no limit: TODO implement pager mechanism?• WoRMS: maximum of 50 records by default, allows paging by using an offset parameter (see below for more details) DONE (pager implemented)• PESI: maximum of 50 records by default, no paging possible so far (see below for more details) DONE (pager implemented)• PLAZI: DONE (pager implemented)• GBIF: DONE			
WoRMS & PESI			
Both are developed and maintained by VLIZ.			
The PESI and WoRMS web services have a couple of operations where the amount of returned records is limited to 50.			
For WoRMS some operations have an optional 'offset' parameter by which paging through the full result set is possible:			
<ul style="list-style-type: none">• getAphiaRecords• getAphiaRecordsByVernacular• getAphiaRecordsByDate• getAphiaChildrenByID			
That's very useful. The PESI service in contrast is missing the ability to retrieve data beyond that 50 records limit. It would strongly increase the utility of the PESI service if you could implement the 'offset' parameter also for PESI. I guess this should not cause too much work, since this already has been done for WoRMS and both implementations seems to be very similar (at least from looking at the SOAP API).			
I asked Bart Vanhoorne from VLIZ for implementation of the 'offset' parameter in PESI (2016-10-21):			
It would strongly increase the utility of the PESI service if you could implement the 'offset' parameter also for PESI. I guess this should not cause too much work, since this already has been done for WoRMS and both implementations seems to be very similar (at least from looking at the SOAP API).			

His Answer:

OK, we'll put this on our todo list.
I would say this will be done end 2016, begin 2017

This is implemented now!!!!

Notify:

- Scott Chamberlain scott@ropensci.org once this is done!

Related issues:

Related to EDIT - feature request #8356: Algae Names Index - implementation	In Progress
Related to EDIT - task #8363: PESI SOAP service seems to return duplicate rec...	Workforme
Related to EDIT - feature request #8360: paging results of /classification/{t...	Workforme

Associated revisions

Revision e34c2dfb - 07/09/2019 05:03 PM - Andreas Kohlbecker

ref #6151 updating PESIClient to latest SOAP api

Revision a6d66952 - 07/09/2019 05:24 PM - Andreas Kohlbecker

ref #6151 pager parameters in UtisController

Revision 11ca8ad3 - 07/09/2019 05:30 PM - Andreas Kohlbecker

ref #6151 pager parameters TnrMsg

Revision 236f847c - 07/09/2019 05:31 PM - Andreas Kohlbecker

ref #6151 pager parameters in UtisController - 2

Revision 8edd1056 - 07/09/2019 10:44 PM - Andreas Kohlbecker

ref #6151 paging of scientificNameLike and vernacularNameLike search modes in WoRMS client

Revision 6fd7dc60 - 07/09/2019 11:13 PM - Andreas Kohlbecker

ref #6151 paging of scientificNameLike and vernacularNameLike search modes in PESI client

Revision 54c8f6e1 - 07/10/2019 08:44 AM - Andreas Kohlbecker

ref #6151 paging of scientificNameLike and vernacularNameLike search modes in EUNIS client

Revision 1a1fb1f0 - 07/10/2019 08:51 AM - Andreas Kohlbecker

ref #6151 paging of scientificNameLike and vernacularNameLike search modes in PLAZI client

Revision 044553f6 - 07/10/2019 12:46 PM - Andreas Kohlbecker

ref #6151 paging of scientificNameLike and vernacularNameLike search modes in BgbmEditClient client

Revision 03b34296 - 07/10/2019 03:06 PM - Andreas Kohlbecker

ref #6151 paging of scientificNameLike and vernacularNameLike search modes in Phycobank client by basing on new common abstract cdm server client which is now also used for the BgbmEditClient

Revision 2e351aa3 - 07/11/2019 02:35 PM - Andreas Kohlbecker

ref #6151 implementing pager in GBIF client and fixing problems with like search

History

#1 - 10/21/2016 01:30 PM - Andreas Kohlbecker

- Status changed from New to In Progress

To save time I decided to change the strategy. I will not implement a general pager mechanism into utis. Not all of the connected services are support paging of the results. The new strategy is to limit the possible amount of returned matched by limiting the minimum search string length:

- [utis-core|34e21a0](#) - testing pager function of worms
- [utis-webappl|bdf44ee](#) - limiting amount of returned results by query length restriction

#2 - 10/21/2016 02:32 PM - Andreas Kohlbecker

- Description updated

#3 - 10/21/2016 02:40 PM - Andreas Kohlbecker

- Due date set to 01/15/2017
- Priority changed from New to Highest
- % Done changed from 0 to 20

For the time until the pager is really implemented I worked on the interim solution by relaxing the max hit setting for the EDIT name catalogue service. This attempt is implemented in [utis-core|4f08022](#). This is however not yet working since the name catalogue service is not yet supporting POST requests. Support for POST requests has been implemented with [cdm|lib|0842c38](#) and will officially become available with the CDM Release 4.4.

This ticket should be continued at the latest in Jan. 2017, setting due date accordingly

#4 - 01/11/2017 11:43 AM - Andreas Kohlbecker

- Description updated

Hi Andreas,

We have now implemented this, and (on top) added a REST service.
<http://www.eu-nomen.eu/portal/webservices.php>

Happy holidays!

Regards,
 Bart

#5 - 07/08/2019 03:46 PM - Andreas Kohlbecker

- Related to feature request #8356: Algae Names Index - implementation added

#6 - 07/09/2019 10:49 PM - Andreas Kohlbecker

- Related to feature request #8360: paging results of /classification/{taxonId}/children webservice added

#7 - 07/09/2019 11:16 PM - Andreas Kohlbecker

the paging mechanism for the WoRMS and PESI clients is implemented.

The PESI SOAP service however seems to return duplicate records:

```
http --timeout 600 :8081/search 'providers==pesi' 'query==Crepis foetida' searchMode==scientificNameLike pageI
ndex==0 pageSize==100 | jq '.query[] | .response[] | .taxon | .url, .taxonName.scientificName '
```

copied to [#8363](#)

#8 - 07/09/2019 11:16 PM - Andreas Kohlbecker

- % Done changed from 20 to 40

#9 - 07/09/2019 11:26 PM - Andreas Kohlbecker

- Description updated
- % Done changed from 40 to 30

#10 - 07/10/2019 09:01 AM - Andreas Kohlbecker

- Description updated

#11 - 07/11/2019 01:04 PM - Andreas Kohlbecker

- Description updated

#12 - 07/11/2019 02:37 PM - Andreas Kohlbecker

- Related to task #8363: PESI SOAP service seems to return duplicate records added

#13 - 07/11/2019 04:28 PM - Andreas Kohlbecker

- Description updated

#14 - 07/11/2019 04:28 PM - Andreas Kohlbecker

- Status changed from In Progress to Resolved

- % Done changed from 30 to 50

#15 - 07/11/2019 04:41 PM - Andreas Kohlbecker

Email sent to Scott Chamberlain

#16 - 07/23/2019 04:06 PM - Andreas Kohlbecker

- Status changed from Resolved to Closed

- % Done changed from 50 to 100

#17 - 07/23/2019 04:07 PM - Andreas Kohlbecker

- Target version set to 280

#18 - 07/23/2019 04:08 PM - Andreas Kohlbecker

- Related to deleted (feature request #8360: paging results of /classification/{taxonId}/children webservice)

#19 - 07/23/2019 04:08 PM - Andreas Kohlbecker

- Related to feature request #8360: paging results of /classification/{taxonId}/children webservice added

#20 - 07/23/2019 04:15 PM - Andreas Kohlbecker

- Tags set to phycobank

#21 - 07/23/2019 04:30 PM - Andreas Kohlbecker

- Category set to utis

#22 - 02/25/2022 04:28 PM - Andreas Müller

- Target version changed from 280 to UTIS 1.3

#23 - 02/25/2022 04:33 PM - Andreas Müller

- Project changed from UTIS to EDIT

- Category changed from utis to utis

- Severity set to normal

#24 - 03/11/2022 03:05 PM - Andreas Müller

- Due date set to 03/11/2022

- Start date set to 03/11/2022

#25 - 03/11/2022 03:18 PM - Andreas Müller

- Due date deleted (03/11/2022)

- Start date deleted (03/11/2022)