

EDIT - task #8363

PESI SOAP service seems to return duplicate records

07/11/2019 02:36 PM - Andreas Kohlbecker

Status:	Worksforme	Start date:	
Priority:	Priority14	Due date:	
Assignee:	Andreas Kohlbecker	% Done:	100%
Category:	utis	Estimated time:	0:00 hour
Target version:			
Severity:	normal		
Description			
see			
<pre>http --timeout 600 :8081/search 'providers==pesi' 'query==Crepis foetida' searchMode==scientificNameLike pageIndex==0 pageSize==100 jq '.query[] .response[] .taxon .url, .taxonName.scientificName '</pre>			
Related issues:			
Related to EDIT - bug #6151: implement pager for search service			Closed

History

#1 - 07/11/2019 02:37 PM - Andreas Kohlbecker

- Related to bug #6151: implement pager for search service added

#2 - 07/11/2019 02:37 PM - Andreas Kohlbecker

- Subject changed from PESI SOAP service however seems to return duplicate records to PESI SOAP service seems to return duplicate records

#3 - 07/11/2019 02:39 PM - Andreas Kohlbecker

Hi Anderas, I added you to this ticket as watcher since this could be an issue wit the data delivery to VLIZ

#4 - 07/15/2019 12:12 PM - Andreas Müller

- Assignee set to Andreas Kohlbecker

can you say which data is duplicated?

Does it appear as duplicate in the eu-nomen portal as well?

#5 - 07/15/2019 12:58 PM - Andreas Kohlbecker

The search for "Crepis foetida%" (`http --timeout 600 https://cybertaxonomy.eu/eu-bon/utis/1.3/search 'providers==pesi' 'query==Crepis foetida' searchMode==scientificNameLike pageIndex==0 pageSize==100 | jq '.query[] | .response[] | .taxon | .url, .taxonName.scientificName'`) returns the folloing records (first line url second line scientific name):

```
"http://www.eu-nomen.eu/portal/taxon.php?GUID=1EC61FFF-E6E4-4F16-98E2-DF3C8179A1C4"
"Crepis foetida L."
"http://www.eu-nomen.eu/portal/taxon.php?GUID=2FF8FD72-DD4A-47DC-B69E-147CABF36B8D"
"Crepis commutata (Spreng.) Greuter"
"http://www.eu-nomen.eu/portal/taxon.php?GUID=47EF102A-BB43-46C2-B900-5AC32BEDEF1D"
"Crepis foetida subsp. glandulosa (C. Presl) Arcang."
"http://www.eu-nomen.eu/portal/taxon.php?GUID=1EC61FFF-E6E4-4F16-98E2-DF3C8179A1C4"
"Crepis foetida L."
"http://www.eu-nomen.eu/portal/taxon.php?GUID=1EC61FFF-E6E4-4F16-98E2-DF3C8179A1C4"
"Crepis foetida L."
"http://www.eu-nomen.eu/portal/taxon.php?GUID=31098C2E-530E-43BA-B9E6-E6AC900F3D6C"
"Crepis foetida subsp. rhoeadifolia (M. Bieb.) Čelak."
"http://www.eu-nomen.eu/portal/taxon.php?GUID=48BD42CA-5AF7-4D8A-857E-80EB507CC408"
"Crepis foetida subsp. sitiaca Rech. f."
"http://www.eu-nomen.eu/portal/taxon.php?GUID=1EC61FFF-E6E4-4F16-98E2-DF3C8179A1C4"
"Crepis foetida L."
"http://www.eu-nomen.eu/portal/taxon.php?GUID=1EC61FFF-E6E4-4F16-98E2-DF3C8179A1C4"
"Crepis foetida L."
"http://www.eu-nomen.eu/portal/taxon.php?GUID=195D8D60-F41C-4966-93B8-BCAF51170445"
```

"Crepis foetida subsp. zacinthia (DC.) Hayek"
"http://www.eu-nomen.eu/portal/taxon.php?GUID=1EC61FFF-E6E4-4F16-98E2-DF3C8179A1C4"
"Crepis foetida L."
"http://www.eu-nomen.eu/portal/taxon.php?GUID=BF508EEE-AD2C-4675-9814-0BC81B645B69"
"Crepis foetida subsp. maritima (Boiss.) Hayek"
"http://www.eu-nomen.eu/portal/taxon.php?GUID=1EC61FFF-E6E4-4F16-98E2-DF3C8179A1C4"
"Crepis foetida L."
"http://www.eu-nomen.eu/portal/taxon.php?GUID=A187E9F6-95BD-44B1-B169-651B131DEC50"
"Crepis foetida subsp. vulgaris (Bisch.) Babç."

Crepis foetida L. with the UUID 1EC61FFF-E6E4-4F16-98E2-DF3C8179A1C4 occurs multiple times.

This exactly is what the Pesi SOAP service is returning. The PESI Portal does not seem to have these duplicates

#6 - 07/15/2019 12:58 PM - Andreas Kohlbecker

- Assignee changed from Andreas Kohlbecker to Andreas Müller

#7 - 07/15/2019 01:10 PM - Andreas Müller

- Assignee changed from Andreas Müller to Andreas Kohlbecker

Why did you assign this to me. As it does not seem to be data issue but a service issue I can't see what I can do.

#8 - 07/15/2019 01:16 PM - Andreas Kohlbecker

If you are sure that this is not a data or data submission issue, I will delegate this problem to Bart Vanhoorn.

#9 - 07/15/2019 01:17 PM - Andreas Müller

As the uuid seems to be the same I don't think we have data duplicates

#10 - 07/15/2019 01:55 PM - Andreas Kohlbecker

- Status changed from New to In Progress

- Priority changed from New to Priority14

- % Done changed from 0 to 10

I send an email to Bart

#11 - 07/23/2019 04:30 PM - Andreas Kohlbecker

- Category set to utis

#12 - 07/24/2019 10:27 AM - Andreas Müller

Interesting that the UTIS project has a category utis ;-)

#13 - 07/24/2019 12:02 PM - Andreas Kohlbecker

I introduced this while I was working on the ticket overview page for Phycobank:

<https://dev.e-taxonomy.eu/redmine/projects/algenregistrierung/wiki/TicketOverview>

Now that I split the tickets for the index and the registration system the category is no longer needed.

#14 - 07/24/2019 12:18 PM - Andreas Kohlbecker

- Status changed from In Progress to Worksforme

- % Done changed from 10 to 100

It turned out that the duplicates are not in the PESI response but that they are an effect of an error in the downstream processing in the UTIS PESI Client. This issue has been fixed.

#15 - 02/25/2022 05:15 PM - Andreas Müller

- Project changed from UTIS to EDIT

- Category changed from utis to utis

- Severity set to normal