PhycoBank - task #7808

Further name duplicates

10/08/2018 08:49 AM - Andreas Kohlbecker

Status:	New		**Start date:**	
Priority:	Highest		**Due date:**	
Assignee:	Wolf-Henning Kusber		**% Done:**	0%
Category:	Datacleaning		**Estimated time:**	0:00 hour
Target version:	Data cleaning phase 2			

Description

copied from [#7420#note-19](#)

existing IAPT data sometimes do have different ranks then the imported data. Therefore the names/taxa are not deduplicated. Example: Cryptophyceae is Division in IAPT but Phylum in Frey + Worms. This needs to be sorted out before the final import.

```
SELECT tn2.uuid, tn2.titleCache, r.titleCache
from TaxonName tn2 join DefinedTermBase r on tn2.rank_id = r.id
where tn2.titleCache in (Select titleCache from (select tn.titleCache as titleCache, count(*) as n
FROM TaxonName tn
GROUP BY tn.titleCache
Having n > 1 ) AS TMP_TBL) order by tn2.titleCache
```

----> [duplicate-names.ods](#)

```
SELECT tn2.uuid, tn2.nameCache, r.titleCache
from TaxonName tn2 join DefinedTermBase r on tn2.rank_id = r.id
where tn2.titleCache in (Select nameCache from (select tn.nameCache as nameCache, count(*) as n
FROM TaxonName tn
GROUP BY tn.nameCache
Having n > 1 ) AS TMP_TBL) order by tn2.nameCache
```

----> [duplicate-names-2.ods](#)

Related issues:

Related to PhycoBank - task #7748: Genus name duplicates, Genus without refer...	**Resolved**
Copied from EDIT - task #7420: Import for higher taxon graph for phycobank	**Closed**

History

#1 - 10/08/2018 08:49 AM - Andreas Kohlbecker

- Copied from task #7748: Genus name duplicates, Genus without reference added

#2 - 10/08/2018 08:53 AM - Andreas Kohlbecker

- Description updated

#3 - 10/08/2018 08:54 AM - Andreas Kohlbecker

- File duplicate-names.ods added

- File duplicate-names-2.ods added

- Description updated

#4 - 10/08/2018 08:55 AM - Andreas Kohlbecker

- Copied from task #7420: Import for higher taxon graph for phycobank added

#5 - 10/08/2018 08:55 AM - Andreas Kohlbecker

- Copied from deleted (task #7748: Genus name duplicates, Genus without reference)

#6 - 10/08/2018 08:55 AM - Andreas Kohlbecker

- Related to task #7748: Genus name duplicates, Genus without reference added

#7 - 05/28/2019 11:46 AM - Andreas Kohlbecker

- Category changed from Import to Datacleaning

#8 - 09/03/2019 04:43 PM - Andreas Kohlbecker

- Target version changed from Registry released to Data cleaning phase 2

Files

duplicate-names-2.ods	18.6 KB	10/08/2018	Andreas Kohlbecker
duplicate-names.ods	32.7 KB	10/08/2018	Andreas Kohlbecker