

EDIT - bug #6938

Tasks before a new Euro+Med import

09/08/2017 06:13 PM - Andreas Müller

Status:	Resolved	Start date:	
Priority:	Priority14	Due date:	
Assignee:	Andreas Müller	% Done:	90%
Category:	data	Estimated time:	2:00 hours
Target version:	Euro+Med Migration	Found in Version:	
Severity:	normal		
Description			
Not yet complete			
see also #5082			
====			
check that the import is adapted such that only referenced authors and teams are imported			
===			
check that import is run with a maven compiled code (aspectJ switched on)			
===			
run [s_em_MoveCommonNames] to move common names to structured common names			
===			
remove duplicate parent relationships (see also http://wiki.bgbm.org/bdinotes/index.php/EuroMed#Data_publication)			
<pre>SELECT * FROM (SELECT COUNT(pt2.RIdentifier) AS HowMany, pt.PTNameFk, pt.PTRefFk , n.FullNameCache FROM PTaxon pt INNER JOIN RelPTaxon rel ON pt.PTNameFk = rel.PTNameFk1 AND pt.PTRefFk = rel.PTRefFk1 INNER JOIN PTaxon pt2 ON rel.PTNameFk2 = pt2.PTNameFk AND rel.PTRefFk2 = pt2.PTRefFk INNER JOIN Name n ON pt.PTNameFk = n.NameID WHERE (rel.RelQualifierFk = 1) AND (pt.PTRefFk % 100000 = 0) AND (pt.PTRefFk >= 7000000) GROUP BY pt.RIdentifier, pt.PTNameFk, pt.PTRefFk, n.FullNameCache) AS DuplicateParentRelation WHERE HowMany > 1 ORDER BY HowMany DESC; SELECT n2.fullNameCache, rel.* FROM PTaxon pt INNER JOIN RelPTaxon rel ON pt.PTNameFk = rel.PTNameFk1 AND pt.PTRefFk = rel.PTRefFk1 INNER JOIN Name n2 ON rel.PTNameFk2 = n2.NameId WHERE rel.PTNameFk1 = xxx AND rel.PTRefFk1 = 7xxx</pre>			
====			
remove selfreferencing relations			
<pre>DELETE RelName WHERE NameFk1 = NameFk2 AND RelNameQualifierFk <> 62; DELETE RelPTaxon WHERE PTNameFk1 = PTNameFk2 and PRefFk1 = PRefFk2</pre>			
====			
Taxa without parents (still needs improvements with publish flag and other issues like valueless taxa)			
<pre>SELECT n.fullNameCache, pt.* FROM PTaxon pt</pre>			

```

INNER JOIN Name n ON pt.PTNameFk = n.NameId
WHERE pt.RIdentifier NOT IN
(SELECT p.RIdentifier
FROM PTaxon p
INNER JOIN RelPTaxon rel ON p.PTNameFk = rel.PTNameFk1 AND p.PTRefFk = rel.PTRefFk1
WHERE (rel.RelQualifierFk = 1)
)
AND (pt.PTRefFk % 100000 = 0) AND (pt.PTRefFk >= 7000000) AND pt.StatusFk NOT IN (2,6, 3,4)

```

====

Run Queries in Y:\BDI\PESI\sh\EuroPlusMed\Credits_Publish\Update_before_publication.sql for data integrity

====

Run Commands in Y:\BDI\PESI\sh\EuroPlusMed\Credits_Publish\SetPublishAndCredits.sql to set the publish flag correctly and to set last scrutiny

===

The export starting point xxx requires that the Treeindex is set correctly in BM.

For this run

```

UPDATE PTaxon SET TreeIndex=NULL;
UPDATE PTaxon SET TreeIndex=dbo.f_core_getTreeIndex(RIdentifier) WHERE (StatusFk IN (1,5))

```

This differs a bit from http://wiki.bgbm.org/bdinotes/index.php/EuroMed#Data_publication as there unpublished taxa do not need a treeindex.

We may also run

```

UPDATE PTaxon SET IndexNameString=dbo.f_core_TreeIndex2NameString(TreeIndex)

```

but not really necessary for the export

===

check for corrupt hierarchies:

```

SELECT n.fullNameCache, n.NameCache, pt.*
FROM PTaxon pt INNER JOIN Name n ON n.NameId = pt.PTNameFk
WHERE pt.TreeIndex NOT LIKE '533921-%' AND pt.TreeIndex NOT LIKE '556619-%'
-- AND pt.StatusFk NOT IN (2,3,4,6)
AND pt.PTRefFk >= 7000000
AND (PublishFlag = 1 OR pt.PTRefFk <> 8000000)
AND (pt.PTRefFk % 100000 = 0)
AND DoubtfulFlag <> 'i'
-- AND IsExcludedMarker = 0
ORDER BY PTRefFk,Treeindex, StatusFK
;
SELECT n.fullNameCache, n.NameCache, pt.*
FROM PTaxon pt INNER JOIN Name n ON n.NameId = pt.PTNameFk
WHERE ptRefFk <> 500000 AND (pt.PTNameFk IN (
SELECT pt.PTNameFk
FROM PTaxon pt
WHERE pt.TreeIndex NOT LIKE '533921-%' AND pt.TreeIndex NOT LIKE '556619-%'
-- AND pt.StatusFk NOT IN (2,3,4,6)
AND pt.PTRefFk >= 7000000
AND (PublishFlag = 1 OR pt.PTRefFk <> 8000000)
AND (pt.PTRefFk % 100000 = 0)
AND DoubtfulFlag <> 'i'
) OR n.nameCache IN ( SELECT n.NameCache
FROM PTaxon pt INNER JOIN Name n ON n.NameId = pt.PTNameFk
WHERE pt.TreeIndex NOT LIKE '533921-%' AND pt.TreeIndex NOT LIKE '556619-%'
-- AND pt.StatusFk NOT IN (2,3,4,6)

```

```

AND pt.PTRefFk >= 7000000
AND (PublishFlag = 1 OR pt.PTRefFk <> 8000000)
AND (pt.PTRefFk % 100000 = 0)
AND DoubtfulFlag <> 'i')
)
ORDER BY n.nameCache, PTNameFK, PTRefFk

```

===

remove inactive taxon relationships not needed anymore

```

SELECT      RelPTaxonId, PTNameFk1, PTRefFk1, PTNameFk2, PTRefFk2, RelQualifierFk, RelRefFk,
Created_When, Updated_When, Created_Who, Updated_Who, Notes, Provisional
FROM        RelPTaxon
WHERE       (RelPTaxonId IN
            (SELECT      rel.RelPTaxonId
             FROM        RelPTaxon AS rel LEFT OUTER JOIN
                        PTaxon AS pt1 ON rel.PTNameFk1 = pt1.
PTNameFk AND rel.PTRefFk1 = pt1.PTRefFk LEFT OUTER JOIN
                        PTaxon AS pt2 ON rel.PTNameFk2 = pt2.
PTNameFk AND rel.PTRefFk2 = pt2.PTRefFk
             WHERE      (rel.RelQualifierFk = - 99) AND (pt1.DoubtfulFlag <>
'i' OR
                        pt1.DoubtfulFlag IS NULL) AND (pt2.
DoubtfulFlag <> 'i' OR
                        pt2.DoubtfulFlag IS NULL)))

```

===

Related issues:

Related to EDIT - task #7341: [Check] CommonName import issues

Resolved

History

#1 - 03/09/2018 09:37 PM - Andreas Müller

- Tags set to euro+med

#2 - 04/05/2018 04:04 PM - Andreas Müller

- Related to task #7341: [Check] CommonName import issues added

#3 - 05/09/2018 05:41 PM - Andreas Müller

- Estimated time set to 2:00 h

#4 - 05/25/2018 12:38 AM - Andreas Müller

- Description updated

#5 - 05/25/2018 12:39 AM - Andreas Müller

- Description updated

- Status changed from New to In Progress

#6 - 05/25/2018 12:43 AM - Andreas Müller

- Description updated

#7 - 05/25/2018 12:45 AM - Andreas Müller

- Description updated

#8 - 05/25/2018 12:51 AM - Andreas Müller

- Description updated

#9 - 05/25/2018 09:06 AM - Andreas Müller

- Description updated

- Priority changed from New to Highest

#10 - 05/25/2018 09:07 AM - Andreas Müller

- Description updated

#11 - 05/25/2018 09:07 AM - Andreas Müller

- % Done changed from 0 to 20

#13 - 05/28/2018 12:55 PM - Andreas Müller

- Description updated

#14 - 05/28/2018 01:03 PM - Andreas Müller

- Description updated

#15 - 05/28/2018 05:01 PM - Andreas Müller

- Description updated

#18 - 05/29/2018 04:59 PM - Andreas Müller

- Description updated

#19 - 05/29/2018 05:00 PM - Andreas Müller

- Description updated

#20 - 05/30/2018 12:05 PM - Andreas Müller

- Description updated

#21 - 05/30/2018 12:13 PM - Andreas Müller

- Description updated

#22 - 06/02/2018 12:14 PM - Andreas Müller

- % Done changed from 20 to 50

#23 - 08/01/2018 02:17 PM - Andreas Müller

- Description updated

#24 - 09/20/2018 03:54 PM - Andreas Müller

- Target version changed from Euro+Med Portal Release to Euro+Med Migration

#25 - 09/29/2018 01:00 AM - Andreas Müller

- Priority changed from Highest to Priority14

#26 - 09/29/2018 12:06 PM - Andreas Müller

- Description updated

#28 - 10/08/2018 05:23 PM - Andreas Müller

- Description updated

#29 - 05/13/2019 10:02 AM - Andreas Müller

- Description updated

#30 - 05/18/2019 08:42 PM - Andreas Müller

- Description updated

#31 - 05/19/2019 09:49 PM - Andreas Müller

- Description updated

#32 - 05/19/2019 10:44 PM - Andreas Müller

- Description updated

#33 - 05/19/2019 11:08 PM - Andreas Müller

- *Status changed from In Progress to Resolved*

- *% Done changed from 50 to 90*

should be resolved, I did run all the steps (1 mail to ERS is still open, about corrupt hierarchies for Polygonum etc.)