

Edit - task #7435

ERS: Handle also published in correctly

05/24/2018 10:58 AM - Andreas Müller

Status:	In Progress	Start date:	05/24/2018
Priority:	Priority13	Due date:	
Assignee:	Andreas Müller	% Done:	70%
Category:	data	Estimated time:	3.00 hours
Target version:	Euro+Med Migration		
Severity:	normal		

Description

Some names do have a second nomenclatural reference (or at least a second reference). This is usually handled as name fact (NameFactCategoryFk = 1).

But sometimes this information is also to be found in some of the notes fields (detail Notes, taxonNotes, nameNotes, ...).

- This information needs to be checked, and in case it is correct moved to the name fact to be imported correctly if not already duplicated there.
- Also the import needs to be adapted to import name facts of cat=1 as either
 - original source of type "2nd nom ref" or "also published in" (both do not exist yet) or "nomenclatural reference" or
 - name fact of type "also published in" (does not exist yet and name facts will be removed soon)
 - extension of type "also published in" (does not exist yet)

Discuss with ERS what the exact semantic is in E+M for this category.

Also published in name facts:

```
SELECT
n.NameCache, n.FullNameCache, pt.PTRefFk, fact.NameFactId, fact.PTNameFk, fact.NameFact, n.notes,
pt.Notes,
fact.NameFactCategoryFk, fact.NameFactRefFk, fact.DoubtfulFlag, fact.PublishFlag,
fact.Created_When, fact.Updated_When, fact.Created_Who, fact.Updated_Who, fact.Notes, fact.NameFactRefDetail
FROM NameFact fact
INNER JOIN Name AS n ON fact.PTNameFk = n.NameId
LEFT OUTER JOIN PTaxon pt ON pt.PTNameFk = fact.PTNameFk AND pt.PTRefFK NOT IN (500000)
WHERE (fact.NameFactCategoryFk = 1)
ORDER BY pt.PTRefFk, n.FullNameCache
```

====

Excel:

- AlsoPublishedIn.xlsx

Related issues:

Related to Edit - task #7796: [CHECK] Handle Bibliography Name Facts correctly	In Progress	09/28/2018
Related to Edit - feature request #7653: [DISCUSS] Delete NameFacts => Result...	Closed	08/14/2018
Related to Edit - feature request #7797: Second nomenclatural reference	New	09/28/2018
Duplicated by Edit - bug #7456: E+M: NameFacts von ILDIS Daten	Duplicate	06/07/2018
Blocks Edit - feature request #7437: Handle also published in data portal	New	05/25/2018

History

#1 - 05/24/2018 11:00 AM - Andreas Müller

- Target version changed from Euro+Med post migration to Euro+Med Migration

#2 - 05/25/2018 09:26 AM - Andreas Müller

- Description updated

- Estimated time changed from 3.00 h to 4.00 h

#3 - 05/25/2018 09:35 AM - Andreas Müller

- Description updated

#4 - 05/25/2018 09:44 AM - Andreas Müller

- Description updated

#5 - 05/25/2018 10:04 AM - Andreas Müller

- Description updated

- % Done changed from 0 to 20

Cleaned up Name.notes duplicates (1 exception send to ERS)

#6 - 05/25/2018 10:25 AM - Andreas Müller

Hallo Eckhard,

in der angefügten Datei befinden sich alle Name Facts vom Type „also published in“. Sind nicht so viele (ca. 120).

Allerdings haben sie sehr unterschiedliche Formate.

Daher einige Fragen:

- Können wir bei denen mit Klammern und & die Klammer und das & löschen?
- Gibt es einen semantischen Unterschied zwischen {& xxx} und {= xxx} oder können wir letztere entsprechend obigem Punkt behandeln?
- Wie oben, aber für {or xxxx}
- Wird die n.v. Information noch benötigt oder kann die weg? (hatten wir schon mal in anderem Zusammenhang, müsste nachschauen, was wir da entschieden haben)
- Soll das „emend.“ Bestehen bleiben? Was bedeutet es in diesem Zusammenhang?
- Sollten wir bei unvollständigen Referenzen diese evtl. vervollständigen (z.B. {& ed. 1812: 35. 1812})? Die unvollständige Referenz setzt eine Formatierung voraus direkt hinter der nom. ref.. Es ist unklar, ob diese immer so erfolgen kann.
- Was bedeutete nochmal „in schedis“?
- Sollten wir die nom. nud. Fälle nicht mit einem separaten Namen abhandeln, der dann die entsprechende Referenz bekommt?
- [or?] Pereira?, Anais Fac. Sci. Porto 17: 233. 1932 ist etwas unsystematisch. Sollten wir da nicht lieber das doubtful flag setzen?
- Sollte eine Statusangabe wie bei {Soják in Cas. Nár. Muz., Odd. Prír. 150: 139. 1982, comb. inval.} nicht wo ganz anders stehen?

Weitere Fragen:

- Bei einigen Facts ist das publish Flag gesetzt, bei einigen nicht. Weißt du, ob das irgendeine Auswirkung hat? Ich habe im alten Portal auf Anhieb keine Seite gefunden, wo diese Information veröffentlicht wurde.
- Soll die Information in Zukunft in irgendeiner Form im neuen Portal veröffentlicht werden?
- Wie ist generell die Semantik? Handelt es sich immer/i.d.R. um eine 2. nomenklatorische Referenz oder ist es einfach irgendeine nicht näher definierte weitere wichtige Publikation?

Kannst du die Daten säubern bzw. mir Regeln an die Hand geben, wie ich sie säubern soll?

Vermutlich gibt es noch deutlich mehr solche Fälle, die in Notes Feldern versteckt sind und noch nicht in den Name Facts. Da werde ich gleich noch nach suchen.

Viele Grüße,
Andreas M.

#7 - 05/25/2018 10:26 AM - Andreas Müller

- Description updated

#8 - 05/25/2018 10:27 AM - Andreas Müller

- Subject changed from *Handle second nomenclatural references correctly* to *Handle also published in correctly*

#9 - 05/25/2018 11:03 AM - Andreas Müller

anbei eine Datei mit „also published in“ Informationen, die sich noch in Name Notes verstecken.

In der ersten Tabelle „Clear“ befinden sich die Fälle, die ziemlich sicher „also published in“ sind.

Könntest du das nochmal bestätigen?

Diese sollten wir in die Name Facts rüberziehen, wobei die Klammern möglichst weggelassen werden sollten (siehe meine letzte Mail).

Könntest du das machen, oder, wenn du es vor deiner Abreise nicht mehr schaffst, mir Bescheid geben, ob alle Fälle ok sind, und ich sie rüberziehen soll? Und welches Format.

In der Tabelle Candidates befinden sich mögliche weitere Fälle, allerdings auch viele Notes, die nicht mit „also published in“ zu tun haben.

Könntest du

- Die Fälle markieren, die auf „also published in“ zutreffen
- Sie entweder selber rüberziehen oder mir die Tabelle schicken, so dass ich es machen kann, sofern du nicht mehr dazu kommst?

Excel: AlsoPublishedInNameNotes.xlsx

#10 - 05/25/2018 11:04 AM - Andreas Müller

- % Done changed from 20 to 40

- Estimated time changed from 4.00 h to 3.00 h

#11 - 05/25/2018 11:06 AM - Andreas Müller

- Precedes feature request #7437: *Handle also published in data portal* added

#12 - 05/25/2018 11:07 AM - Andreas Müller

- Precedes deleted (feature request #7437: *Handle also published in data portal*)

#13 - 05/25/2018 11:07 AM - Andreas Müller

- Blocks feature request #7437: *Handle also published in data portal* added

#14 - 05/25/2018 11:27 AM - Andreas Müller

hier eine überarbeitete Version der Excel Datei mit einer 3. Tabelle, die einige Fälle enthält, bei denen im Detail Note die entsprechende „also published in“ Information steht.

Bitte auch in dieser, überprüfen, ob das so stimmt, und ggf. in die NameFacts ziehen.

#16 - 09/20/2018 03:57 PM - Andreas Müller

- Priority changed from New to Highest

#17 - 09/28/2018 06:22 PM - Andreas Müller

- Description updated

#18 - 09/28/2018 08:02 PM - Andreas Müller

- Related to task #7796: [CHECK] Handle Bibliography Name Facts correctly added

#19 - 09/28/2018 08:51 PM - Andreas Müller

... achso, und es gibt auch noch einige, die nicht mit „&“ oder „=“ oder „or“ anfangen, wie die meisten anderen. Ist das korrekt oder soll das noch angepasst werden?

FullNameCache	NameFact
Hieracium prediliense var. prassbergense Hayek	an jam Sched. Fl. Stiriac. 21-22: [in schedis] No. 1099. 1911?
Hieracium pseudozetlandicum Roffey	descr. in J. Bot. 63: 316. 1925
Garhadiolus hedypnois Jaub. & Spach	emend. Boissier, Fl. Orient. 3: 723. 1875
Artemisia alpina Willd.	emend. M. Bieb., Fl. Taur.-Caucas. 2: 297. 1808
Hieracium subortum Gus. Schneid.	Gus. Schneid. in Deutsche Bot. Monatsschr. 6: 122. 1888, nom. nud.
Hieracium vulgatum var. hepaticum Lindeb.	in Bot. Not. 1878: 184. 1878
Hieracium melanotrichum Reut.	in Compt.-Rend. Trav. Soc. Hallér. 2: 25. 1854
Hieracium xanthoprasinophyes Gottschl.	Zahn in Ascherson & Graebner, Syn. Mitteleur. Fl. 12(3): 418. 1939, descr. germ., ex

```

SELECT
  n.NameCache, n.FullNameCache, fact.NameFact, n.NomRefFk, pt.PTRefFk, fact.NameFactId, fact.PTNameFk, n.notes,
  pt.Notes,
  fact.NameFactCategoryFk, fact.NameFactRefFk, fact.DoubtfulFlag, fact.PublishFlag,
  fact.Created_When, fact.Updated_When, fact.Created_Who, fact.Updated_Who, fact.Notes, fact.NameFactRefDet
ail,
  pt.PublishFlag
FROM NameFact fact
  INNER JOIN Name AS n ON fact.PTNameFk = n.NameId
  LEFT OUTER JOIN PTaxon pt ON pt.PTNameFk = fact.PTNameFk AND pt.PTRefFK NOT IN (500000)
WHERE (fact.NameFactCategoryFk = 1) AND (fact.NameFact not like '%&' AND fact.NameFact not like '%=' AND
  fact.NameFact not like '%or%' AND fact.NameFact not like '%n.v.%' )
ORDER BY fact.NameFact, pt.PTRefFk, n.FullNameCache;

```

Die folgenden Namen haben "Also published in" Namefacts, die entweder leer (erster) sind, oder ein n.v. bzw. ein des. inval. haben. Das kann natürlich richtig sein, aber insbesondere bei n.v. bin ich mir nicht sicher, ob die nicht entweder falsch geparsed oder falsch kopiert wurden, da sie davor stehen und mit Semikolon getrennt sind. Kannst du ja mal überprüfen. Für den Import ist es allerdings nicht wirklich ein Problem, da ich sie erstmal sowieso nur als Text importiere.

Reichardia gaditana subsp. lusitanica (Cout.) P. Silva & Bacelar

Adenostyles hybrida Guss.	& (Vill.) DC., Prodr. 5: 204. 1836, des. inval. (nom. prov.)
Hieracium ausserdorferi Oborny	& [in schedis] No. 3373; etiam: Zahn in Koch, Syn. Deut. Schweiz. Fl., ed. 3: 1821. 1901
(Sep.), des. inval. sine descr. specif.	
Santolina pinnata Viv.	& Donn, Hortus Cantabrig.: 93. 1796, des. inval. (nom. nud.)
Hieracium bifidum subsp. hemilonchodes Zahn	& in Ascherson & Graebner, Syn. Mitteleur. Fl. 12(2): 666. 1935, des. inval. (descr. germ.)
Hieracium carinthiostiriicum J. Vetter & Zahn	& in Ascherson & Graebner, Syn. Mitteleur. Fl. 12(3): 698. 1938, des. inval. (descr. germ.)
Hieracium levicaule subsp. sublonchodes Murr & Zahn	& Murr & Zahn in Ascherson & Graebner, Syn. Mitteleur. Fl. 12(2): 709. 1935, des. inval. (descr. germ.)
Reynoutria sachalinensis (F. Schmidt) Nakai	& Nakai, Rep. Veg. Ooryongto: 18. 1919, des. inval. (nom. nud.)
Reynoutria sachalinensis (F. Schmidt) Nakai	& Nakai, Rep. Veg. Ooryongto: 18. 1919, des. inval. (nom. nud.)
Hieracium solacolui Prodan	& Nyár. in Savulescu, Fl. Rep. Pop. Române 10: 723. 1965, des. inval. (sine design. typi)
Hieracium liptoviense subsp. coinotokum O. Behr & al.	& O. Behr & al. in Ascherson & Graebner, Syn. Mitteleur. Fl. 12(3): 193. 1936, des. inval. (descr. germ.)
Doronicum clusii var. villosum Beck	& Tausch in Flora 11: 178. 1828, des. inval.
Hieracium liptoviense subsp. amaurosphaera J. Vetter & Zahn & Vetter & Zahn	& in Ascherson & Graebner, Syn. Mitteleur. Fl. 12(3): 194. 1936, des. inval. (descr. germ.)
Picris auriculata Sch. Bip.	{n.v., & in Bot. Zeitung (Berlin) 23: 226. 1865}
Lactuca viminea subsp. chondrilliflora (Boreau) St.-Lag.	n.v.; & Bonnier, Fl. Ill. France 6: 79. 1923
Hieracium rhaeticum Froel.	n.v.; & in Rchb., Fl. Germ. Excurs.: 260. 1831-1832
Achillea millefolium subsp. setacea (Waldst. & Kit.) Čelak.	n.v.; & Prodr. Fl. Böhmen: 228. 1871
Achillea millefolium var. alpestris Wimm. & al.	n.v.; & Wimm. & Grab., Fl. Siles. 2(2): 169. 1829
Crepis triasii (Cambess.) Fr.	n.v.; or Nyman, Syll. Fl. Eur.: 49. 1854-1855

```

SELECT
  n.NameCache, n.FullNameCache, fact.NameFact, n.NomRefFk, pt.PTRefFk, fact.NameFactId, fact.PTNameFk, n.notes,
  pt.Notes,
  fact.NameFactCategoryFk, fact.NameFactRefFk, fact.DoubtfulFlag, fact.PublishFlag,
  fact.Created_When, fact.Updated_When, fact.Created_Who, fact.Updated_Who, fact.Notes, fact.NameFactRefDet
ail,
  pt.PublishFlag
FROM NameFact fact
  INNER JOIN Name AS n ON fact.PTNameFk = n.NameId
  LEFT OUTER JOIN PTaxon pt ON pt.PTNameFk = fact.PTNameFk AND pt.PTRefFK NOT IN (500000)
WHERE (fact.NameFactCategoryFk = 1) AND (fact.NameFact like '%des. inv%' OR fact.NameFact like '' OR
  fact.NameFact like '%n.v.%' )
ORDER BY fact.NameFact, pt.PTRefFk, n.FullNameCache;

```

#20 - 09/28/2018 11:12 PM - Andreas Müller

- Related to bug #7456: E+M: NameFacts von ILDIS Daten added

#21 - 09/28/2018 11:14 PM - Andreas Müller

- Related to feature request #7653: [DISCUSS] Delete NameFacts => Result: no added

#22 - 09/28/2018 11:49 PM - Andreas Müller

- Related to feature request #7797: Second nomenclatural reference added

#23 - 09/28/2018 11:57 PM - Andreas Müller

- Related to deleted (bug #7456: E+M: NameFacts von ILDIS Daten)

#24 - 09/28/2018 11:58 PM - Andreas Müller

- Duplicated by bug #7456: E+M: NameFacts von ILDIS Daten added

#25 - 09/29/2018 12:47 AM - Andreas Müller

- Subject changed from Handle also published in correctly to ERS: Handle also published in correctly
- Status changed from New to In Progress
- Priority changed from Highest to Priority13
- % Done changed from 40 to 70

Only the remaining 2 comments need to be handled by ERS (see mail). Maybe also some name notes in CDM.

The issue itself is blocked or will be handled by [#7797](#).

#26 - 09/29/2018 12:48 AM - Andreas Müller

We keep it as a name fact for now.