

EDIT - bug #8858

Ordering of distribution status 'undefined' is not correct in E+M

02/24/2020 01:44 PM - Andreas Müller

Status:	Closed	Start date:	
Priority:	Highest	Due date:	
Assignee:	Andreas Müller	% Done:	100%
Category:	cdmlib	Estimated time:	0:00 hour
Target version:	Release 5.21	Found in Version:	
Severity:	normal		
Description			
AM			
<p>guter Hinweis. Grundsätzlich hat das nicht mit der Transmission-Engine zu tun, sondern mit der on-the-fly Berechnung, welcher Status angezeigt werden soll (erstere ist primär für die Aggregation der Daten von einem Rang zum anderen zuständig). Das Problem hier scheint zu sein, dass der Status „undefined“ nicht aus dem sortierten Standardvokabular kommt sondern aus einem separaten „User Defined“ Vokabular, da der Term für E+M extra angelegt wurde. Derzeit werden Terme aus verschiedenen Vokabularen deterministisch aber undefiniert sortiert. Daher läuft das hier wohl noch falsch. Datenmodellseitig sind inzwischen die Grundlagen gelegt, dass das auch anders geht, so dass wir uns da mal hinsetzen sollten um das umzustellen. Das sollte vor dem offiziellen Release passieren.</p> <p>Ich lege ein Ticket an für den E+M Fall, auch wenn das sowieso auf der Liste steht.</p> <p>===</p> <p>ERS:</p> <p>ich frage mich gerade, weshalb bei einer Verbreitungsangabe mit unterschiedlichen Status für dasselbe Areal der "schlechtere" Status für die Ausgabe im Portal ausgewählt wird, und nicht derjenige, der laut unserer festgelegten Hierarchie den anderen "überschreiben" soll. Das sieht doch so aus, als sei der Tansmission Engine eben noch nicht gelaufen (dazu gab es unterschiedliche Aussagen, als ich letzten Donnerstag mit Walter darüber gesprochen hatte).</p> <p>Beispiel:</p> <p>Für Centaurea chalcidicaea und Griechenland https://euoplusmed.org/cdm_dataportal/taxon/60d8dc3b-ac65-4346-9869-85898518fc2a</p> <p>wird im Portal der Status "undefined" mit einer uralten Quelle ausgegeben.</p> <p>Es gibt aber einen record für dasselbe Areal mit dem Status "native" und vier neueren Quellen, der stattdessen agezeigt werden sollte (und im BerlinModell auch korrekt ausgegeben wird),</p> <p>Wie kann das im CDM verwirklicht werden? Das ist eine entscheidende Voraussetzung für das Release des Portals.</p>			
Related issues:			
Related to EDIT - feature request #3816: Create "Undefined" status and choose...		Closed	
Related to EDIT - bug #8297: Fix condensed distribution string for E+M		Closed	
Related to EDIT - bug #9477: New ordered terms do not have correct order index		Closed	
Copied to EDIT - feature request #9500: Allow removing certain distribution s...		Feedback	

Associated revisions

Revision b553ef04 - 02/24/2021 05:16 PM - Andreas Müller

ref #8858 ordering of terms from different vocabularies use voc.uuid as vocabulary order attribute (fix ordering of "undefined" status in E+M)

Revision 8b7c16f0 - 02/24/2021 10:00 PM - Andreas Müller

ref #3816, ref #8858 add "undefined" to PAT vocabulary

Revision 4dccf2d6 - 03/02/2021 01:36 PM - Andreas Müller

fix #8858 Improve presence/absence handling for PresenceAbsenceTerm.compareTo

History

#2 - 02/24/2021 05:20 PM - Andreas Müller

- Status changed from New to Resolved
- Target version changed from Euro+Med Portal Release to Release 5.21
- % Done changed from 0 to 40

Fixed by adding ordering by vocabulary uuid to term ordering. This is a preliminary solution until ordering via ordered termtree is not yet possible in the webservice.

We should open a new ticket to fully fix this. Also further data cleaning should be done on distributions with status undefined.

#3 - 02/24/2021 05:38 PM - Andreas Müller

- Related to feature request #3816: Create "Undefined" status and choose it for NULL status added

#4 - 02/24/2021 05:38 PM - Andreas Müller

Another solution is implementing [#3816](#)

#5 - 02/24/2021 10:40 PM - Andreas Müller

- Related to bug #8297: Fix condensed distribution string for E+M added

#6 - 02/25/2021 10:27 AM - Andreas Müller

- Related to bug #9477: New ordered terms do not have correct order index added

#7 - 02/27/2021 08:25 AM - Andreas Müller

- Status changed from Resolved to Closed
- % Done changed from 40 to 100

Andreas Müller wrote:

Another solution is implementing [#3816](#)

This solution finally has been implemented. The status priority seems to work now.

#8 - 03/01/2021 10:59 AM - Andreas Müller

- Related to bug #4409: fallback areas are not adressable in the shape file added

#9 - 03/01/2021 11:01 AM - Andreas Müller

- Related to deleted (bug #4409: fallback areas are not adressable in the shape file)

#11 - 03/02/2021 01:34 PM - Andreas Müller

- Status changed from Closed to Resolved

The reason why the first fix didn't work is that PresenceAbsenceTerm had it's own compareTo method, including an isPresent/isAbsent check. Present terms were generally preferred to absence terms which for "undefined" did not work as "undefined" is "present" as isAbsent is not a three-state variable and the state is simply not known. Therefore even after including "undefined" into the CDM vocabulary the comparison did not work for comparing undefined with absent terms.

This is fixed now the way that the presence/absence check is only done if the vocabularies differ (the assumption that presence terms should win still seems to be more reasonable then the assumption that the vocabulary with the lower uuid should win).

Reset to review for checking if the result is as expected.
Example data can be found for *Talinum paniculatum*/Madeira.

#12 - 03/02/2021 01:36 PM - Andreas Müller

- % Done changed from 100 to 50

Applied in changeset [cdmlib|4dccb2d61ae938ff7638aa262de483068859fd3b](#).

#13 - 03/02/2021 03:40 PM - Andreas Müller

- Copied to feature request #9500: Allow removing certain distribution status from distribution publication added

#14 - 03/02/2021 03:41 PM - Andreas Müller

- Status changed from Resolved to Closed

- % Done changed from 50 to 100

The new implementation works as expected. The result is for *Talinum paniculatum* on test is "-[] -Md(??M) [aAu(A) pAz(aF aP) aHu -Md]" http://test.e-taxonomy.eu/dataportal/preview/euromed/cdm_dataportal/taxon/3f28aea3-829f-46dc-957a-9991c30050f4

However, further research shows that the "undefined" distributions were not shown at all in the Berlin Model system. This is because they were put into brackets, which means that they are handled as false or unreliable information.

E.g.: <http://ww2.bgbm.org/EuroPlusMed/PTaxonDetailOccurrence.asp?NameId=48608&PRefFk=7300000>

So we probably need functionality to fully remove "undefined" distributions from any publication output (condensed distribution string, textual distribution and map).

This will be handled in a new ticket: [#9500](#)