

BIEN 3 - Task #917

Task # 928 (New): switch to new TNRS setup

TNRS: Instructions for new version with TPL

05/13/2014 02:32 PM - Martha Narro

Status:	New	Start date:	05/13/2014
Priority:	Normal	Due date:	
Assignee:	Aaron Marcuse-Kubitza	% Done:	80%
Category:		Estimated time:	0.00 hour
Target version:			
Activity type:	Coding/analysis		
Description			
From Brad:			
I've given some thought to the TPL matter. The algorithm isn't hard, but Aaron will have to do the sorting himself.			
<ol style="list-style-type: none">1. Make sure sources are selected in the following order: GCC, TPL, Tropicos, USDA2. When downloading names, do NOT sort by source (i.e. don't limit results to just the best match when sorted by source)3. Download all results (not just best matches)3a. fix anomaly where there were multiple Selected names for some input names (to avoid breaking constraints)3b. reimplement the parsed rank columns for the all matches strategy, which does not have a single scrubbed name per input name to parse see taxon_match_derived_columns .3c. create table and algorithm to store a selected best match for each input name4. Apply the usual TNRS sort order (see the README .¹) to the matches for a give name. Aaron, Brad said there are no additional steps to apply here (in step 4). Just proceed to step 5. (- Martha) Brad now says that actually the TNRS sort order is incorrect because of the Constrain by Source bug .5. If the best match (indicated by Selected=TRUE) has source=Tropicos and acceptance=accepted AND another match is available where source<is not equal to>Tropicos and acceptance=synonym, use the latter name (we don't need this until after the names are scrubbed (Martha))6. All other cases, use the best match as flagged			
That should filter out most Tropicos nomenclatural synonyms incorrectly labeled accepted. I can unpack #4 for Aaron when the time arrives.			
Brad			
¹ note that edit_distance = (1 - specific_epithet_score)*greatest_length			

History

#1 - 05/15/2014 09:43 AM - Martha Narro

Aaron, for now (next couple of months, this must be done on the development TNRS web app since TPL is only on the dev app. Martha will send you the url.

#2 - 05/20/2014 01:52 PM - Aaron Marcuse-Kubitza

- Description updated

#3 - 05/20/2014 04:28 PM - Aaron Marcuse-Kubitza

- Parent task set to #928

#4 - 05/27/2014 11:15 PM - Aaron Marcuse-Kubitza

- Description updated

- % Done changed from 0 to 20

#5 - 05/28/2014 10:29 AM - Martha Narro

- Description updated

#6 - 06/02/2014 04:50 AM - Aaron Marcuse-Kubitza

- Description updated

- % Done changed from 20 to 40

#7 - 06/02/2014 03:59 PM - Aaron Marcuse-Kubitza

- Description updated

#8 - 06/02/2014 04:16 PM - Aaron Marcuse-Kubitza

- Description updated

#9 - 06/02/2014 04:17 PM - Aaron Marcuse-Kubitza

- Description updated

#10 - 06/02/2014 04:27 PM - Aaron Marcuse-Kubitza

- Description updated

#11 - 06/04/2014 06:00 PM - Martha Narro

- File 48_test_names.txt added

- File 48_test_name_tnrs_results.xlsx added

The taxon names Brad sent for testing the rescrubbing are now attached.

#12 - 06/18/2014 12:49 PM - Martha Narro

- Description updated

#13 - 06/18/2014 01:53 PM - Aaron Marcuse-Kubitza

- % Done changed from 40 to 60

all the taxon names have now been rescrubbed

#14 - 06/24/2014 03:19 PM - Aaron Marcuse-Kubitza

- Description updated

#15 - 06/24/2014 03:56 PM - Aaron Marcuse-Kubitza

- Description updated

#16 - 06/24/2014 03:56 PM - Aaron Marcuse-Kubitza

- Description updated

#17 - 06/24/2014 03:58 PM - Aaron Marcuse-Kubitza

- Description updated

#18 - 06/25/2014 11:20 AM - Aaron Marcuse-Kubitza

- Description updated

#19 - 06/25/2014 11:44 AM - Aaron Marcuse-Kubitza

- Description updated

#20 - 06/25/2014 11:46 AM - Aaron Marcuse-Kubitza

- Description updated

#21 - 06/25/2014 11:52 AM - Aaron Marcuse-Kubitza

- Description updated

#22 - 06/25/2014 07:48 PM - Aaron Marcuse-Kubitza

- Description updated

multiple Selected names bug fixed in [r13855](#)

#23 - 06/26/2014 09:19 AM - Aaron Marcuse-Kubitza

- Description updated

#24 - 06/26/2014 09:23 AM - Aaron Marcuse-Kubitza

testing Martha's watcher notifications...

#25 - 06/26/2014 09:23 AM - Aaron Marcuse-Kubitza

- Description updated

#26 - 06/26/2014 09:25 AM - Aaron Marcuse-Kubitza

testing Martha's watcher notifications again after e-mail address fix...

#27 - 06/26/2014 09:37 AM - Aaron Marcuse-Kubitza

- Description updated

#28 - 06/26/2014 06:49 PM - Aaron Marcuse-Kubitza

- Description updated

#29 - 06/26/2014 06:51 PM - Aaron Marcuse-Kubitza

- Description updated

- % Done changed from 60 to 70

#30 - 06/27/2014 03:15 PM - Aaron Marcuse-Kubitza

test update now that watcher notifications have been fixed

#31 - 07/03/2014 12:46 AM - Aaron Marcuse-Kubitza

- % Done changed from 70 to 80

- Description updated

#32 - 07/03/2014 03:47 AM - Aaron Marcuse-Kubitza

- Description updated

a single "-" creates strikethrough formatting, so you have to use "--" instead

#33 - 07/10/2014 01:00 AM - Aaron Marcuse-Kubitza

- Description updated

- % Done changed from 80 to 70

#34 - 07/18/2014 02:25 PM - Aaron Marcuse-Kubitza

- Description updated

#35 - 07/18/2014 04:34 PM - Aaron Marcuse-Kubitza

- Description updated

#36 - 08/04/2014 07:39 AM - Aaron Marcuse-Kubitza

- Description updated

- % Done changed from 70 to 80

#37 - 08/05/2014 03:42 PM - Aaron Marcuse-Kubitza

- Description updated

#38 - 08/14/2014 04:42 PM - Aaron Marcuse-Kubitza

- Description updated

- % Done changed from 80 to 90

#39 - 08/25/2014 05:19 PM - Aaron Marcuse-Kubitza

- *Description updated*

- *% Done changed from 90 to 80*

#40 - 04/28/2017 04:38 AM - Aaron Marcuse-Kubitza

- *Description updated*

Files

48_test_names.txt	1.22 KB	06/05/2014	Martha Narro
48_test_name_tnrs_results.xlsx	42.2 KB	06/05/2014	Martha Narro