

BIEN 3 - Task #816

re-run TNRS on mis-scrubbed names

11/18/2013 10:50 PM - Aaron Marcuse-Kubitza

Status:	New	Start date:	11/18/2013
Priority:	Normal	Due date:	
Assignee:	Aaron Marcuse-Kubitza	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:			
Activity type:			

Description

- temporary workaround for names with an accepted name:
use Accepted_name_family when Name_matched_accepted_family is not available
- this will fix the 300,000 no-opinion names¹ scrubbed during the 6 days that TNRS was broken², when it wasn't populating the critical field Name_matched_accepted_family

¹[293816](#)

```
SELECT COUNT(*) FROM "TNRS".tnrs
WHERE
    "Name_matched_accepted_family" IS NULL
AND "Family_matched" IS NOT NULL
AND "Accepted_name_family" IS NULL
-- where we are using the matched name instead of the accepted name
AND "Name_matched_rank" != 'family' -- this is also the case for family-only matches
Total query runtime: 5718 ms.
```

² 2013-5-23..28

```
SELECT min(batch), max(batch) FROM "TNRS".tnrs
WHERE
    "Name_matched_accepted_family" IS NULL
AND "Family_matched" IS NOT NULL
AND "Accepted_name_family" IS NULL
-- where we are using the matched name instead of the accepted name
AND "Name_matched_rank" != 'family' -- this is also the case for family-only matches
Total query runtime: 3928 ms.
"2013-05-23 15:48:28.970796-07", "2013-05-28 08:36:52.569103-07"
```

History

#1 - 11/18/2013 10:56 PM - Aaron Marcuse-Kubitza

- Description updated

#2 - 11/18/2013 11:05 PM - Aaron Marcuse-Kubitza

- Description updated

#3 - 11/18/2013 11:10 PM - Aaron Marcuse-Kubitza

- Description updated

#4 - 11/19/2013 01:16 PM - Aaron Marcuse-Kubitza

- Description updated

#5 - 01/09/2014 09:44 AM - Aaron Marcuse-Kubitza

- Description updated