

PhycoBank - task #6014

Use cdmlib validation framework for algae names

07/25/2016 03:11 PM - Andreas Kohlbecker

Status:	New	Start date:	07/25/2016
Priority:	Priority10	Due date:	
Assignee:		% Done:	0%
Category:	Validation	Estimated time:	0.00 hour
Target version:			
Description			
Das Validierungsframework des CDM soll in der AlgenRegistrierung für die Validierung der Algennamen bezüglich der Regeln des ICN verwendet werden.			
Allerdings wären, um es für Algen nutzbar zu machen, auch die zahlreichen Sonderregelungen für Algennamen zu bedenken. Für Pflanzen müssen die meisten Regeln auch noch implementiert werden. Bisher ist das nur für einige Grundregeln passiert, prototypisch.			
Subtasks:			
task # 6015: Assemble ICN rules special for algae names			New
task # 6092: Illustration in algal descriptions			New
feature request # 6016: ICN rules in cdm validation framework implemented			New
Related issues:			
Related to PhycoBank - task #6505: Check if type names have never been used a...		New	03/14/2017
Related to PhycoBank - task #6506: Check if a genus has been used before in b...		New	03/14/2017

History

#1 - 07/25/2016 03:12 PM - Andreas Kohlbecker

Das Validierungsframework für die Algenregistrierung zu nutzen ist absolut sinnvoll. Dieser Prozessschritt sollte auch in den Workflows berücksichtigt werden.

Ein wichtiger Aspekt der dabei beachtet werden muss ist, dass die Validierung nicht synchron mit der Speicherung eines Namen passiert sondern ein asynchroner Schritt ist.

Das Validierungsergebnis liegt also erst nach dem Speichern vor.

Das stimmt nur teilweise. Auf jeden Fall stimmt, dass die Validierung bei der jetzigen Implementierung erst beim Speichern passiert, das Ergebnis also frühestens nach dem Speichern vorhanden ist. Das ist im Normalfall auch so gewollt, da das kein invasiver Eingriff sondern letztlich nur eine Hilfestellung sein soll, um die Daten zu säubern. In einigen Fällen, wird es aber auch nach dem Speichern noch eine Zeit dauern, bis das Ergebnis, zu sehen ist. Dies betrifft komplexere Zusammenhänge, die beim einfachen Speichern nicht so einfach abgeprüft werden können. Für die Algen sollte das aber nur bedingt relevant sein, da es sich hier ja primär um Namensdaten handelt, also eher simple Strukturen. Vielleicht wird es relevant bei einigen Regeln, bei denen der Typen etc. auch mit reinspielen.

Interessant ist hier aber, dass das Framework und insbesondere auch die Regeln aber genauso gut auch anders verwendet werden können, nämlich zur ad-hoc Validierung. Wir verwenden das bislang nur nicht so. Bei der Registrierung könnte ich mir vorstellen, dass das von Interesse ist, um unsaubere Daten möglichst früh zu erkennen und gar nicht erst unsauber ins System zu lassen. Wobei ihr ja sicherlich auch so etwas wie eine Staging-Area etc. plant, so dass auch unsaubere Daten erstmal gespeichert werden können, bevor sie dann gesäubert und ins endgültige System übertragen werden.

Viele Grüße,
Andreas M.

#2 - 03/15/2017 09:41 AM - Andreas Kohlbecker

- *Category set to Validation*

#3 - 03/15/2017 09:42 AM - Andreas Kohlbecker

- *Related to task #6505: Check if type names have never been used as type name before added*

#4 - 03/15/2017 09:43 AM - Andreas Kohlbecker

- *Related to task #6506: Check if a genus has been used before in botany and zoology added*