memo

|  |
| --- |
| Taskforce MAS |

|  |  |
| --- | --- |
| To: | , |
| From: | Pim van Avesaath – IHM |
| CC: |  |
| Date: | Date |
| Re: | Verkenning noodzaak en mogelijkheden voor de mapping van de soortenregisters WoRMS en TWN list. |
| Comments: |  |
| Status: | draft-v 0.7 |

WoRMS: World Register of Marine Species - The aim of a World Register of Marine Species (WoRMS) is to provide an authoritative and comprehensive list of names of marine organisms, including information on synonymy. While the highest priority goes to valid names, other names in use are included so that this register can serve as a guide to interpret taxonomic literature. <https://www.marinespecies.org/about.php>

TWN: Taxa Waterbeheer Nederland - De TWN lijst geeft een opsomming van alle aquatische organismen in Nederland en is onderdeel van de Aquo-standaard. De lijst wordt onderhouden door Rijkswaterstaat en is in opdracht van Informatie Huis Water (IHW) ontwikkeld. <https://www.ecosys.nl/twn-lijst/> , <https://twnlist.aquadesk.nl/> (account vereist)

Breder kader: mapping van de semantische standaarden AQUO en SeaDataNet/BODC

De validatie van soortnamen van beide standaarden is gebaseerd op verschillende soortenregisters: AQUO gebruikt de TWN lijst, SeaDataNet/BODC gebruikt WoRMS. Voor een goede mapping van de standaarden is ook een mapping tussen de soortenregisters nodig.

Willen Stolte heeft hiervoor reeds verkennende gesprekken gevoerd met specialisten van RWS-CIV. Hieruit is gebleken de koppeling mogelijk is door te matchen op wetenschappelijke namen van taxa en de autoriteit waar de determinatie van het taxon op is gebaseerd (auteursnamen, jaartal) zie tabel 1. Dit is nu al mogelijk door gebruik te maken van de Match Taxa Tool van WoRMS (<https://www.marinespecies.org/aphia.php?p=match> )

Om inzicht te krijgen in de haalbaarheid en efficiëntie van deze methode voor de koppeling is in een experiment uitgevoerd waarbij de taxa van de TWN lijst handmating ( ‘quick and dirty’) zijn gematcht met behulp van de Match Taxa Tool van WoRMS.

Ten tijde van dit experiment (download 15-02-2022) stonden er 27196 taxa vermeld in de TWN lijst, waarvan er 8526 taxa een match vertoonde met WoRMS. Van de resterende 18670 taxa (die niet voorkwamen in de WoRMS lijst) wordt (gemakshalve) aangenomen dat het geen mariene soorten zijn (land/zoetwater taxa), zie Table 2. Het is aan te nemen dat er nog mariene taxa in de TWN lijst ontbreken. Dit zal gedurende het gebruik van de mapping blijken. Voor het opnemen van nieuwe soorten (in beide registers) dient structureel overlegd opgezet te worden met de beheerders van beide registers (WoRMS: VLIZ; TWN: RWS-CIV). Enkele taxa veroorzaakten en time-out van de service en verwijderd uit de analyse.

De analyse van de efficiëntie van de matching verder gediversifieerd worden op basis van het attribuut ’match type’ (Table 2):

* Van de gemeenschappelijke taxa (8526) kan 92% exact gematcht worden op basis van de combinatie wetenschappelijke naam en auteur(s)-jaartal.
* De resterende 8% bevat onduidelijkheden of fouten die additionele aandacht nodig hebben.
  + Zo zijn er koppelingen gemaakt tussen taxa met 1 of meerdere afwijkende tekens (near 1 – 3, phonetic).
  + Van 158 taxa dient er nog handmatig een keuze gemaakt te worden in de koppeling tussen de twee lijsten (match type: ambiguous) vanwege te grote discrepanties in taxonnaam en/of auteursgegevens/jaartal.
  + Dan zijn er ook matches tussen taxa die in WoRMS in quarantaine zijn geplaats of reeds verwijderd zijn.
  + Het betreft voornamelijk problemen met matches van Diatomeeën en fytoplankton (Figure 1, Table 3)

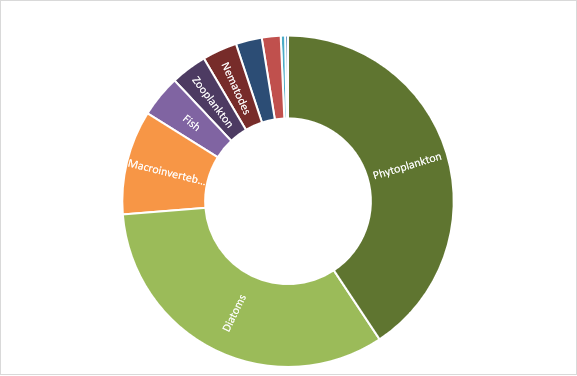


Figure 1. Overview of TWN taxa (summed by TWN taxa types) with an incomplete match with WoRMS

Technisch gezien lijkt de koppeling tussen WoRMS en TWN goed mogelijk te zijn, en met een hoge opbrengst (92 %). Er zijn verschillende methodes beschikbaar om de koppeling tussen de TWN lijst en Worms met behulp van de WORMS tools te automatiseren ( SOAP en REST API’s, zie <https://www.marinespecies.org/aphia.php?p=webservice> ).

Er zal een procedure nodig zijn, waarbij de beheerders van beide registers betrokken worden, om de discrepanties aan te pakken. Waarschijnlijk ook specialisten van moeilijke groepen.

Er schijnen al overlegstructuren te zijn tussen WTN en WoRMS. Het is verstandig om te proberen hierop aan te haken.

Dit experiment betreft slechts de koppeling van de soortenregisters TWN en WoRMS. De integratie met fylogenie is hier nog niet bij betrokken, maar wel een wens. Waarschijnlijk zal dit een ander inzicht brengen in de mogelijkheden en opties voor de koppeling van taxonomie en fylogenie.

Table 1. Column names (attributes) of the TWN and WoRMS output van de matching taxa tables

|  |  |  |
| --- | --- | --- |
| **TWN** |  | **WoRMS matching - output** |
| taxontype |  | AphiaID |
| taxonname |  | Match type |
| author |  | LSID |
| taxongroup |  | TSN |
| taxonlevel |  | Qualitystatus |
| parentname |  | Taxon status |
| refername |  | ScientificName |
| literature |  | Authority |
| localname |  | AphiaID\_accepted |
| date |  | ScientificName\_accepted |
| status |  | Authority\_accepted |
|  |  | Kingdom |
|  |  | Phylum |
|  |  | Class |
|  |  | Order |
|  |  | Family |
|  |  | Genus |
|  |  | Subgenus |
|  |  | Species |
|  |  | Subspecies |
|  |  | isMarine |
|  |  | isBrackish |
|  |  | isFresh |
|  |  | isTerrestrial |
|  |  | Citation |

Table . Result of the taxa match of the TWN with WoRMS using the WoRMS Taxa match tool

|  |  |  |
| --- | --- | --- |
| Match type | Total | % of Matched taxa |
| **exact**- all characters match exactly | 7818 | 92 |
| **exact\_subgenus** - an exact match, but including the subgenus | 25 | 0 |
| **ambiguous** - taxon name or authority | 158 | 2 |
| **near\_1** - perfect match, except for one character. This is a quite reliable match | 68 | 1 |
| **near\_2** - good match, except for two characters. This needs an extra check | 108 | 1 |
| **phonetic**- sounds similar as, despite minor differences in spelling (soundex algorithm) | 157 | 2 |
| **near\_3** - good match, except for three characters. This definitely needs an extra check | 12 | 0 |
| **match\_quarantine** - match with a name that is currently in quarantine. Any name that has been used in the literature should in principle not be quarantined. So best to contact the WoRMS DMT about this | 105 | 1 |
| **match\_deleted** - this is a match with a name that has been deleted and no alternative is available. Please contact the WoRMS DMT when you come across this. | 75 | 1 |
| **Total matched** | **8526** |  |
| **TWN taxa not matched** | **18670** |  |
| **Grand Total** | **27196** |  |

Table 3. Overview of TWN taxa (summed by taxa types) with an incomplete match with WoRMS

|  |  |
| --- | --- |
| taxon type | Total |
| Amphibia | 2 |
| Birds | 13 |
| Diatoms | 234 |
| Fish | 29 |
| Macroalgae | 3 |
| Macroinvertebrates | 72 |
| Macrophytes | 18 |
| Nematodes | 24 |
| Phytoplankton | 288 |
| Zooplankton | 25 |
| Grand Total | 708 |