

AQUARES

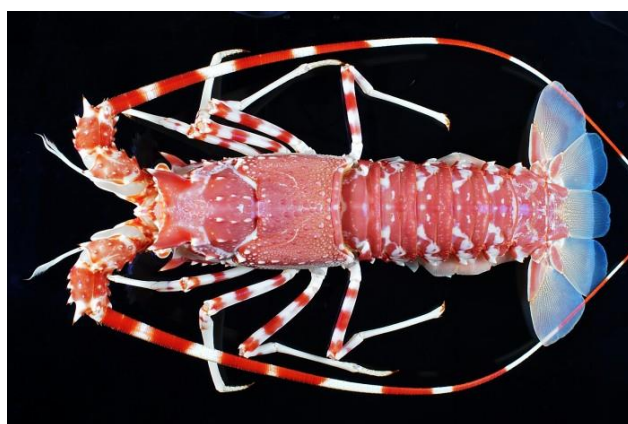
Uitwisseling van Data en Services tussen Aquatische Soortenregisters

DUUR
01/12/2013 - 29/02/2016

BUDGET
316.708 €

PROJECT BESCHRIJVING

Het gebruik van soortnamen is alomtegenwoordig in een breed scala van wetenschappelijke-, milieubeheers- en beleidsdomeinen. Door specialisten beheerde taxonomische databases en zoekfuncties om deze gegevens te bevragen zijn daarom essentieel voor het waarborgen van de kwaliteit van biologische data, gaande van het verzamelen en ontwikkelen van data tot het verbeteren en beheren ervan. Soortinformatiesystemen voor de opvolging van de status en trends van de biodiversiteit en deze relevant voor het beleid – Natura 2000-soorten, commerciële, invasieve exoten en schadelijke soorten – hebben voordeel bij dergelijke toepassingen en databases van hoge kwaliteit die de interoperabiliteit van de gegevens garanderen.



Het World Register of Marine Species (WoRMS), het Register of Antarctic Marine Species (RAMS) en de Freshwater Animal Diversity Assessment (FADA) databank zijn drie belangrijke Global Species Directories (GSD) die in België gehost worden. Deze gegevenscollecties bestaan uit gezaghebbende taxonomische lijsten, samengesteld door internationale experts en dragen bij aan initiatieven zoals Catalogue of Life (CoL) en LifeWatch, en mogelijk aan de Pan-European Species directories Infrastructure (PESI) en andere regionale en nationale soortenlijsten. De meeste van deze initiatieven rekenen op een breed scala aan bijdragen van specialisten aan onafhankelijke checklists en vereisen uitgebreide interacties met een groot netwerk van deskundigen. Gezien de mogelijke overlap in taxonomische specialisten en de complexe aard van de gegevens, is het uitwisselen van kennis en gegevens tussen deze initiatieven zeer gunstig voor alle betrokken partijen.

De belangrijkste doelstelling van dit project is dan ook om de interoperabiliteit en de publieke beschikbaarheid van deze aquatische soortendatabanken te waarborgen en te verbeteren door de ontwikkeling van een reeks van web services. Dergelijke web services kunnen de automatische uitwisseling van gegevens tussen WoRMS, RAMS en FADA garanderen, maar stellen de gegevens ook beschikbaar voor gebruik in andere initiatieven en toepassingen zoals Encyclopedia of Life (EoL), Catalogue of Life (CoL), Global Biodiversity Information Facility (GBIF) en e-Science initiatieven zoals het Biodiversity Virtual e-Laboratory (BioVeL) project en LifeWatch.

Om de kwaliteit van de gegevens die via deze web services worden blootgesteld te waarborgen, streven wij ernaar om de procedures voor het importeren en uitwisselen van gegevens in de partner-databases te verbeteren. In het kader hiervan zullen we ook een interface ontwikkelen om het invoeren en aanvullen van soortverspreidingsgegevens te vergemakkelijken. Deze procedures en tools zullen worden getest en gebruikt tijdens een praktische workshop met taxonomische experten. Om hun betrokkenheid te stimuleren en de vrije en open publicatie van hun gegevens aan te moedigen, zullen wij een dienst ontwikkelen voor het genereren van een checklist artikel. Dit kan worden gepubliceerd in een wetenschappelijk tijdschrift en biedt een eenvoudigere oplossing voor het correct citeren en traceren van citaten van de data.

AQUARES

Bovenop de web services van de individuele databanken, zullen we een gezamenlijke data cache creëren waarop we een scala van tools voor kwaliteitscontrole van soortgerelateerde gegevens zullen implementeren. Een centrale tool, die zowel voor interne kwaliteitscontrole van taxongegevens als voor de curatie van verspreidingsgegevens of andere soortgerelateerde gegevens relevant is, is de zogenaamde Taxamatch tool. Deze tool maakt gebruik van fuzzy matching algoritmen voor het zoeken van fonetisch overeenstemmende of zeer vergelijkbare wetenschappelijke soortnamen, en zal worden aangepast en verbeterd voor het gebruik in combinatie met de gezamenlijke data cache. Andere instrumenten, die meer specifiek zijn gericht op soortverspreidingsgegevens, omvatten onder andere een service om observatiegegevens te vergelijken met door deskundigen gecontroleerde verspreidingsarealen, en hulpmiddelen voor het opsporen van technische fouten in de gegevensbestanden (bv. verkeerde datum-formaat, ontbreken van verplichte velden). Hoewel deze voorzieningen zullen draaien als zelfstandige tools, zijn ze zeer relevant voor een breed scala van gebruikers en zouden ze, in het kader van de Europese Biodiversity Virtual e-Laboratory (BioVeL) project, gecombineerd kunnen worden in een data workflow samen met andere web services die beschikbaar zijn in het biodiversiteitsinformaticadomein.

Doorheen dit project zullen we regelmatig overleg organiseren met een breed scala aan potentiële gebruikers om hun noden te documenteren en hun feedback op de ontwikkelde tools en diensten te krijgen. Gegevens van het FP7 BioFresh-project, het Europese Oceaan Biogeographic Information System (EurOBIS) en de Antarctische Biodiversity Information Facility (AntaBIF) zullen worden gebruikt als specifieke test cases ter validering en verbetering van de tools. Verdere tests met gegevens uit biologische collecties en ecologische monitoringgegevens worden overwogen om ervoor te zorgen dat deze diensten van belang zijn voor een brede waaier aan instituten en onderzoekers die met aquatische soortendata werken.



CONTACT INFORMATIE

Coördinator

Aaike DE WEVER

Koninklijk Belgisch Instituut voor Natuurwetenschappen (KBIN)

Operationele Directie Natuurlijk Milieu

aaike.dewever@naturalsciences.be

Partners

Leen VANDEPITTE

Vlaams Instituut voor de Zee (VLIZ)

Datacentrum

leen.vandepitte@vliz.be

Bruno DANIS

Université libre de Bruxelles (ULB)

Marine Biology Lab

bdanis@ulb.ac.be