

## **Samenvatting van het project**

De Europese natuurwetenschappelijke collecties bevatten het grootste en belangrijkste deel van de wetenschappelijke kennis van de structuur, het milieu en de biosfeer van de aarde. De Belgische collecties bestrijken alle domeinen van de natuurgeschiedenis, met inbegrip van zoölogie, plantkunde, geologie, paleontologie en antropologie. Belgische wetenschappelijke instellingen herbergen meer dan 55 miljoen specimens en vormen één van de belangrijkste collecties met betrekking tot natuurwetenschappen wereldwijd.

De partners van het project zijn ook betrokken bij het opzetten van DiSSCo (dissco.eu), de nieuwe ESFRI gerelateerd aan natuurhistorische collecties met als doel open toegang te verzekeren tot Europese natuurhistorische collecties en daarbij de gebruikersgemeenschap te verbreden en innovatieve oplossingen te vinden door het gebruik van natuurwetenschappelijke gerelateerde data.

De snelle evolutie op het gebied van programmering en gegevensuitwisseling tussen verschillende systemen en toestellen (desktops, laptops, tablets, smartphones, enz.), samen met de exponentiële ontwikkeling van webtoepassingen die toegang bieden tot een grote verscheidenheid van geografisch gecodeerde of aan geografische referenties gerelateerde natuurwetenschappelijke informatie, dwingt tot een aanpak waarbij de onderlinge koppeling van gescheiden modules en gegevens centraal staat.

In het kader van het NaturalHeritage project is onderzocht en geëvalueerd hoe een modulaire architectuur een antwoord kan bieden op de verscheidenheid van gegevens- en collectiebeheersystemen door een gemeenschappelijk onderzoeksinstrument aan te bieden. Deze architectuur is gebaseerd op twee niveaus van interoperabiliteit tussen kleine modules en met internationale authority files. De nieuwe architectuur is gericht op interoperabiliteit binnen het portaal NaturalHeritage.be (tussen de verschillende submodules) en daarbuiten met de gegevens die beschikbaar zijn via gecertificeerde e-bronnen en gegevens die gepubliceerd zijn via internationale aggregatoren. De voorgestelde modulaire organisatie van NaturalHeritage zal ook een betere bruikbaarheid van het systeem mogelijk maken door interne en externe wetenschappers, collectiebeheerders en andere potentiële gebruikers zoals experts uit de verschillende disciplines (Zoölogie, Plantkunde, Geologie, Paleontologie, etc.).

De doelstellingen van het project zijn

- Het bereiken van een beter geïntegreerd digitaal collectiebeheer op Belgisch niveau gebaseerd op de interoperabiliteit van de verschillende datasets en modules met een gemeenschappelijke zoekportaal; Dit betreft de individuele gegevens die reeds gedigitaliseerd zijn in de Collectiebeheersystemen maar ook de beschrijving van de collecties die nog niet gedigitaliseerd zijn (ongeveer 90% van de collecties).

- Voorzien in adequate validatie-instrumenten, met gebruikmaking van webdiensten die voldoen aan internationaal erkende normen voor de digitale curatie van de inhoud als bron (d.w.z. kruiscontrole van wetenschappelijke namen, locatienamen, gebruikmaking van

normen voor de import/export van gegevens, gebruikmaking van normen voor visualisatie-instrumenten, enz.)

- Het (de) gegevensmodel(len) uitbreiden tot nieuwe soorten collecties bij partners die niet op taxonomische gegevens zijn gericht, om een bredere interoperabiliteit en een grotere toepassing te bereiken;

- Voldoen aan gegevensvereisten volgens Europese richtlijnen (INSPIRE & OpenData) en internationale protocollen (Nagoya) en regelgeving (ABS);

- Betrouwbare open toegang bieden tot gegevens van hoge kwaliteit van de verschillende belanghebbenden en gebruikers via een gemeenschappelijk onderzoeksportaal dat een historisch overzicht en een globaal overzicht biedt van biodiversiteit en wetenschappelijke gegevens via een zoekstelsel dat geografisch en/of thematisch sorteert;

- Het promoten van de nieuwe architectuur bij collega's en peers die CETAF gebruiken door de capaciteit van het platform uit te breiden, zowel geografisch (Europa en internationaal) als conceptueel (naar andere verwante disciplines), in het licht van bestaande maatschappelijke uitdagingen (invasieve soorten, gezondheidsveiligheid, enz.).

De methodologie die door het project werd aangenomen, consolideerde en verbreedde de samenwerking, integratie en interoperabiliteit tussen Belgische instellingen en ontwikkelde de technische hub voor de Belgische DiSSCo-partners. Er werd voorrang gegeven aan Open Source technologieën die reeds gebruikt worden door ICT-partners om inspanningen te bundelen en volledige interoperabiliteit te produceren tussen de Belgische partners, maar ook met grotere en bredere internationale systemen zoals het GBIF portaal en die compatibel zullen zijn met het DiSSCo Elvis systeem.

Op basis van deze aanpak ontwikkelde het project een nieuw database-ecosysteem en een zoekportaal, zodat er een interoperabiliteit is tussen de 3 collectiepartners (RBINS, KMMA en APM).

Het project breidde het bestaande DaRWIn-collectiebeheersysteem uit tot nieuwe categorieën van objecten en bijbehorende data/metadata. Het gegevensmodel produceert gegevens en metadata die compatibel zijn met andere internationale inspanningen op milieugebied of voor de archieven (GBIF, BIOCASE, GEOGASE, OAI-PMH) en voldoet aan de verplichtingen van de lidstaten van de Europese Unie (voornamelijk afgeleid van INSPIRE, Open Data en Nagoya Protocol).

Het project heeft ook nieuwe modules ontwikkeld voor specifieke soorten objecten, zoals afbeeldingen, multimedia, archieven in het kader van het Nagoya-protocol en CITES, bibliografieën, enz.

De interoperabiliteit van al deze data/metadata zal wetenschappers en beleidsmakers helpen om Belgische natuurhistorische collecties te benaderen met behulp van cross-linked en big-data benaderingen en het ontwierp de Belgische data hub voor het nieuwe DiSSCo ESFRI.

Het project ontwikkelde ook een nieuw gemeenschappelijk onderzoeksportaal dat gebruikers in staat stelt 3 niveaus van vragen te gebruiken:

- eenvoudige google-zoekopdracht,
- Wie ?, Wanneer ?, Waar ? Hoe ? als eenvoudige "menselijke" vragen
- of gedetailleerd zoeken met facetten

Het gebruik van UUID maakt het mogelijk altijd een link te leggen naar de oorspronkelijke gegevens. Dit is belangrijk voor het oorspronkelijke eigendom van de gegevens, de traceerbaarheid en het citeren van de bron.

Sommige specifieke collecties zoals herbarium maken massale digitaliseringsprogramma's mogelijk, maar de meeste collecties moeten worden gedigitaliseerd en het globale aandeel van de collectie dat op specimenniveau is gedigitaliseerd, bedraagt niet meer dan 10%. De volledige digitalisering van de specimens is dus een doelstelling op zeer lange termijn. Het is zeer belangrijk om ook informatie te verstrekken op het niveau van de collecties en subcollecties, zelfs als deze nog niet gedigitaliseerd zijn. Daarom heeft het NaturalHeritage consortium een instrument ontwikkeld dat de aanpak en de resultaten van het NH-project hergebruikt en dat het mogelijk maakt gegevens/metada op collectieniveau te verstrekken die informatie geven over de omvang, de diversiteit en de taxonomische en/of geografische dekking van deze collecties.

Het project liet de partners toe om belangrijke ontwikkelingen te realiseren in het kader van de voorbereiding van de Belgische collecties op de DiSSCo-infrastructuur en de meeste van de ontwikkelingen zijn gebaseerd op open source technologieën en beschikbaar op de Github-mappen van de projectpartners voor hergebruik door andere instellingen/collecties in België of in het buitenland.

Niettemin maken de verdeling van bevoegdheden en financieringsbronnen het (zeer) moeilijk om expertise/technologieën over te dragen aan geïnteresseerde partners die niet tot dezelfde institutionele pijler behoren (regionaal of federaal). Het BRAIN-programma bevordert de samenwerking tussen Federale Wetenschappelijke Instellingen (FWI) en regionale instellingen op het gebied van onderzoek, maar er is geen financieringsmogelijkheid voor de overdracht van de technologieën/expertise op de lange termijn. Het infrastructuurprogramma van het FWO laat bijvoorbeeld geen financiering toe van federale instellingen, zelfs niet als onderaannemer, en het federale programma ter ondersteuning van de infrastructuur financiert alleen de FWI terwijl de budgettaire besparingen verplichten tot het concentreren van de activiteiten op exclusieve federale competenties. Een duidelijk/gemakkelijk samenwerkingspad tussen het federale en het regionale niveau moet worden opgezet om de samenwerking(en) tussen FWI en wetenschappelijke instellingen van een ander niveau te versterken.

### **Keywords:**

Zoekportaal, Open Source, Collectiebeheerssysteem, Webservices, DiSSCo ESFRI