

DASA

Digital Animal Sound Archive

Contract - B2/233/P2/DASA

Context

Een grote verscheidenheid aan dieren produceert akoestische signalen of roepen, die in veel gevallen soortspecifiek zijn. Bioakoestische monitoring is daarom een essentieel instrument geworden om de verspreiding, ecologie, het gedrag en de fenologie van organismen te bestuderen die vaak visueel moeilijk waarneembaar zijn (bijv. zeezoogdieren, vleermuizen). Akoestische datasets zijn echter vaak versnipperd over instellingen en bestandsformaten, wat hun hergebruik en integratie voor grootschalige ecologische analyses bemoeilijkt.

Het Digital Animal Sound Archive (DASA)-project werd in februari 2023 gelanceerd om deze uitdagingen aan te pakken door een uniforme infrastructuur te creëren voor de opslag, het beheer en de valorisatie van bioakoestische opnamen die in België zijn verzameld.

Het project, gecoördineerd door het Koninklijk Belgisch Instituut voor Natuurwetenschappen (KBIN) in samenwerking met twee milieu-NGO's (Natuurpunt en Natagora), richtte zich in eerste instantie op vleermuisopnamen (Chiroptera) als *proof of concept*.

Doelstellingen

Het belangrijkste doel van DASA was het centraliseren van verspreide bioakoestische datasets in één enkele FAIR (Findable, Accessible, Interoperable en Reusable)-conforme infrastructuur, die gegevens beschermt voor toekomstig onderzoek en tegelijk een betrouwbare referentieverzameling van gevalideerde soortspecifieke geluiden biedt.

Andere doelstellingen zijn:

- Ontwerpen en implementeren van een robuust databasemodel voor akoestische gegevens en metadata, geïnspireerd op bestaande applicaties zoals ecoSound-web, Tethys en waarnemingen.be.
- Ontwikkelen van een gebruiksvriendelijk webplatform voor dataselectie en validatie.
- Opzetten van duurzame redundante opslagoplossingen die datasets van tientallen terabytes aankunnen.
- Stimuleren van samenwerking tussen professionals en burgerwetenschappers om datavalidatie en hergebruik te bevorderen.
- Bevorderen van interoperabiliteit tussen burgerwetenschapsnetwerken en internationale biodiversiteitsplatforms zoals het Global Biodiversity Information Facility (GBIF).

Methodologie

Er werd een uitgebreid databasemodel opgebouwd, met entiteiten voor project, plaatsing, detectie, voorkomen en determinatie, waardoor zowel menselijke als geautomatiseerde identificatieprocessen mogelijk zijn. Een importprocedure werd ontwikkeld die gestandaardiseerde metadatamodellen combineert met automatische validatiescripts om de gegevenskwaliteit en taxonomische consistentie te waarborgen. De systeemarchitectuur bestaat uit een relationele database (PostgreSQL/PostGIS), een ElasticSearch-index voor geavanceerde zoekopdrachten en een webapplicatie (<https://dasa.naturalsciences.be>) gebouwd met Spring Boot en React, die API's, gefacetteerde zoekfuncties, sonogramvisualisaties en meertalige interfaces biedt.

Project B2/233/P2/DASA - Digital Animal Sound Archive

De opslaginfrastructuur bij het KBIN omvat 20 TB primaire netwerkopslag (NAS) met offsite back-ups en langdurige tape-archivering.

Resultaten

Het DASA-project heeft met succes een functionele digitale infrastructuur opgezet die bijna twee miljoen vleermuisopnamen integreert, waarvan ongeveer één miljoen rechtstreeks aan geluidsbestanden gekoppeld zijn. Voor de overige opnamen worden de metadata veilig bewaard. Het platform ondersteunt zowel menselijke als automatische determinaties, transparant metadata-beheer en collaboratieve validatie. Het sluit volledig aan bij de FAIR-principes en de Europese Open Data-richtlijn, en waarborgt openheid, interoperabiliteit en langdurige toegankelijkheid. Hoewel de eerste focus op vleermuizen lag, maakt het systeemontwerp een eenvoudige uitbreiding naar andere taxa en eco-akoestische domeinen mogelijk.

Aanbevelingen

Om het volledige potentieel van de DASA-ontwikkelingen te benutten, worden volgende aanbevelingen voor de toekomst gedaan:

- Toekomstige ontwikkelingen moeten prioriteit geven aan het vereenvoudigen van de data-indieningsprocessen, zodat bijdragers opnamen en metadata rechtstreeks via de webinterface kunnen uploaden.
- Voortgezette betrokkenheid van burgerwetenschappers en professionele netwerken is essentieel om validatieactiviteiten te onderhouden en de datadekking uit te breiden.
- Alle akoestische gegevens, inclusief veelvoorkomende soorten en achtergrondgeluiden, moeten bewaard blijven, aangezien deze van grote waarde zijn voor langetermijn-ecologische en methodologische studies.
- Een plan voor data-opslaguitbreiding bij het KBIN moet worden uitgerold om de opslag van bijkomende datasets te faciliteren.
- Versterking van de interoperabiliteit met externe classificatiepijplijnen en biodiversiteitsdataplatforms zal de wetenschappelijke impact en duurzaamheid van DASA verder vergroten.
- Er is nog steeds nood aan betere referentiegegevens, namelijk visuele bevestigingen van opgenomen geluiden om het doel van een referentiecollectie volledig te realiseren.

Trefwoorden

Bioakoestiek; Digitale infrastructuur; Gegevensbeheer; Burgerwetenschap; Vleermuizen; Ecologische monitoring; Langetermijnopslag.