DASA
Digital Animal Sound Archive

Contract - B2/233/P2/DASA

Context

A wide variety of animals produce acoustic signals or calls, that are in many cases species-specific Bioacoustic monitoring has therefore become an essential tool to study species distribution, ecology, behaviour and phenology of organisms that are often visually elusive (e.g. marine mammals, bats). However, acoustic datasets are often fragmented across institutions and formats, hindering their reuse and integration for large-scale ecological analysis. The Digital Animal Sound Archive (DASA) project was launched in February 2023 to address these challenges by creating a unified infrastructure for the storage, management, and valorisation of bioacoustic recordings collected in Belgium. The project, coordinated by the Royal Belgian Institute of Natural Sciences (RBINS) in collaboration with two environmental NGOs (Natuurpunt, Natagora) focused primarily on bat (Chiroptera) recordings as a proof of concept.

Objectives

The main objective of DASA was to centralise scattered bioacoustic datasets in a single FAIR (Findable, Accessible, Interoperable and Reusable)-compliant infrastructure that safeguards data for future research while providing a reliable reference collection of validated species-specific sounds. Other objectives are:

- Designing and implementing a robust database model for acoustic data and metadata, informed by existing frameworks such as ecoSound-web, Tethys, and waarnemingen.be.
- Developing a user-friendly web platform for data searching and human validation
- Establishing long-term redundant storage solutions capable of handling datasets in the tens of terabytes.
- Encouraging collaboration between professionals and citizen scientists to enhance data validation and reuse.
- Promoting interoperability between citizen science networks and international biodiversity platforms such as the Global Biodiversity Information Facility (GBIF).

Methodology

A comprehensive database model was built, incorporating entities for project, deployment, detection, occurrence, and determination, enabling both human and automated identification processes. A data import procedure was developed that combines a standardized metadata templates with automated validation scripts ensuring data quality and taxonomic consistency. A system architecture was set-up consisting of a relational database (PostgreSQL/PostGIS), an ElasticSearch index for advanced querying and finally, a web application. The web application (https://dasa.naturalsciences.be) is built with Spring Boot and React, providing APIs, faceted search,

visualization of sonograms, and multilingual user interfaces. Storage infrastructure at RBINS includes 20 TB of primary network attached storage (NAS) with offsite backups and long-term tape archiving.

Results

The DASA project successfully established a functional digital infrastructure integrating nearly two million bat recordings, of which about one million are directly linked to audio files. For the other recordings, the metadata are safeguarded. The platform supports incremental human and machine-based determinations, transparent metadata management, and collaborative validation. It fully aligns with the FAIR principles and the European Open Data Directive, ensuring openness, interoperability, and long-term accessibility. Although initially focused on bats, the system's design allows easy extension to other taxa and eco-acoustic domains.

Recommendations

To continue to make use of the full potential of the DASA developments, following recommendations for the future can be made:

- Future developments should prioritise simplifying data submission workflows, allowing contributors to upload recordings and metadata directly through the web interface rather than relying solely on spreadsheets.
- Continued engagement with citizen scientists and professional networks will be essential to sustain validation activities and expand data coverage.
- All acoustic data, including common species and background noise, should be preserved, as they hold significant value for long-term ecological and methodological studies.
- A data storage acquisition plan will need to be rolled out for RBINS to facilitate the storage of more datasets.
- Strengthening interoperability with external classification pipelines and biodiversity data platforms will further enhance the scientific impact and sustainability of the DASA infrastructure.
- Need for better reference data: visual confirmations of recorded sounds to fully fullfil the reference collection objective.

Keywords

Bioacoustics; Digital infrastructure; Data management; Citizen science; Bats; Ecological monitoring; Long-term storage.