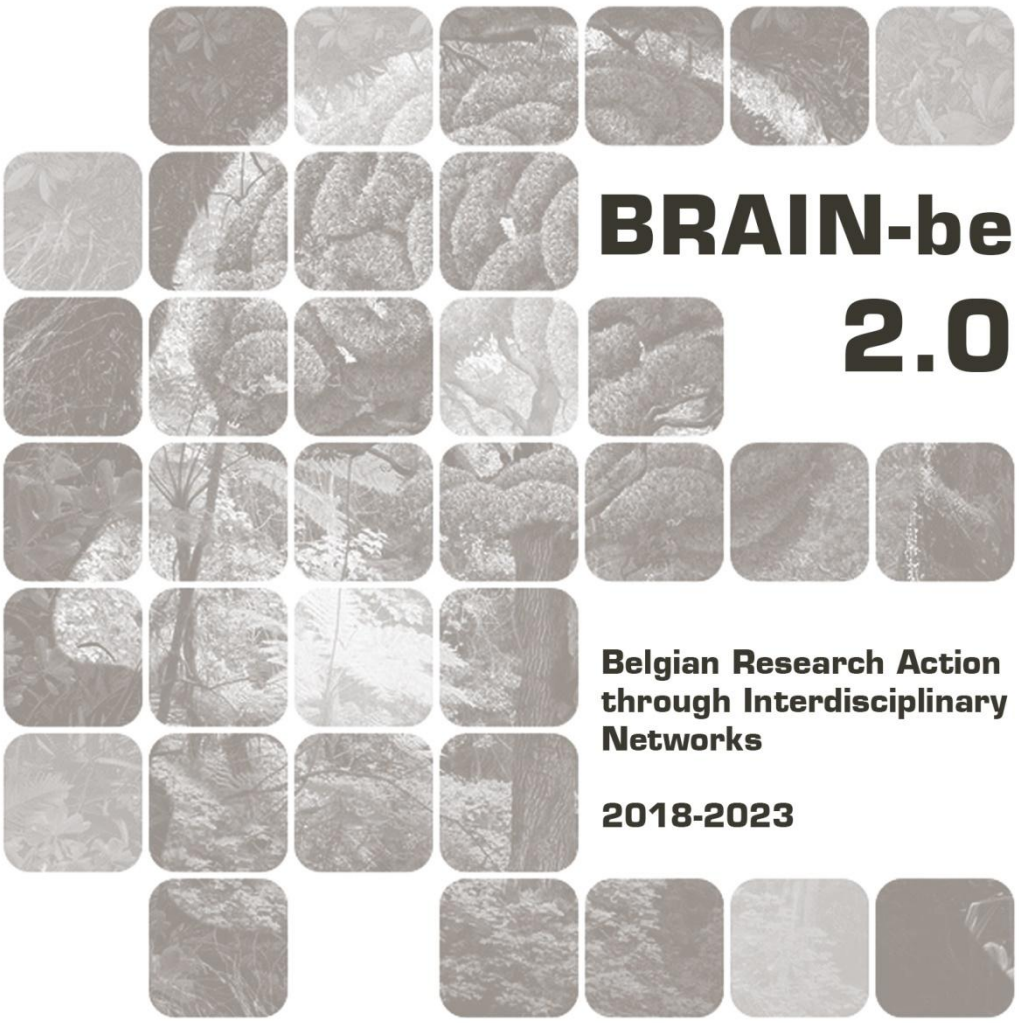

DASA

Digital Animal Sound Archive

Thomas Vandenberghe (RBINS) – Robin Brabant (RBINS) – Yves Laurent (RBINS) – Claire Brabant (Natagora) – Bob Vandendriessche (Regionaal Landschap Houtland & Polders vzw) – Wout Willems (Natuurpunt)

Pillar 2: Heritage science

belspo .be

BRAIN-be 2.0

Belgian Research Action through Interdisciplinary Networks

2018-2023

NETWORK PROJECT

DASA

Contract - B2/233/P2/DASA

FINAL REPORT

PROMOTORS: Thomas Vandenberghe (RBINS)

Robin Brabant (RBINS)

AUTHORS: Thomas Vandenberghe (RBINS)

Robin Brabant (RBINS)

Claire Brabant (Natagora)

Yves Laurent (RBINS)

Wout Willems (Natuurpunt)

natural sciences .be

natagora

natuurpunt

belspo .be

Published in 2025 by the Belgian Science Policy Office
WTCIII
Simon Bolivarlaan 30 bus 7
Boulevard Simon Bolivar 30 bte 7
B-1000 Brussels
Belgium
Tel: +32 (0)2 238 34 11
http://www.belspo.be
http://www.belspo.be/brain-be

Contact person: Georges JAMART
Tel: +32 (0)2 238 36 90

Vandenberghe T., Brabant R., Brabant C., Laurent Y., Willems W. Digital Animal Sound Archive (DASA) Final report. Brussels: Belgian Science Policy Office 2025 – 18 p. (BRAIN-be 2.0 - (Belgian Research Action through Interdisciplinary Networks))

TABLE OF CONTENTS

1. ABSTRACT 5

2. INTRODUCTION 5

3. STATE OF THE ART AND OBJECTIVES 5

4. METHODOLOGY 9
4.1. DATABASE DESIGN AND DATA STORAGE APPROACH 9
4.2. DATABASE AND IMPORT PROCEDURE DEVELOPMENT 11
4.3. DATA ACQUISITION AND VALIDATION 12
4.4. USER INTERFACE DEVELOPMENT 12

5. SCIENTIFIC RESULTS AND RECOMMENDATIONS 13

6. DISSEMINATION AND VALORISATION 14
6.1. WORKSHOPS 14
6.2. PRESENTATIONS 15
6.3. PRESS RELEASES 15
6.4. FOLLOW-UP PROJECTS 15
6.5. FOLLOW-UP COMMITTEE 15
6.6. OTHER VALORISATION EFFORTS 16

7. PUBLICATIONS 16

8. ACKNOWLEDGEMENTS 17

9. REFERENCES 18

1. ABSTRACT

Many animal species produce acoustic signals that are in many cases species-specific and that serve as valuable indicators for studying their distribution, behaviour, and ecology. However, bioacoustic datasets are often fragmented across institutions and formats, limiting their scientific reuse. The Digital Animal Sound Archive (DASA) project, launched in 2023, was developed to centralise and safeguard Belgian bioacoustic recordings, initially focusing on bats as a proof of concept. DASA integrates recordings from research and citizen science initiatives within a unified, FAIR-compliant infrastructure.

The project established a comprehensive metadata model, a scalable storage system for multi-terabyte datasets, and a web application enabling browsing, visualization, and collaborative validation. The database supports incremental automated and human identifications and ensures interoperability with international standards and biodiversity platforms such as GBIF. Nearly two million bat detections, including about one million linked audio files, were successfully incorporated.

By aligning with the FAIR principles and the European Open Data Directive (EU 2019/1024), DASA ensures openness, transparency, and long-term data preservation. Its modular design allows extension to other taxa and acoustic monitoring contexts.

2. INTRODUCTION

A wide variety of animals produce acoustic signals or calls, that are in many cases species-specific (e.g. Berwick et al. 2011, Brabant et al. 2019, Hawkins 1986, Richardson et al. 2013). The use of these animal sounds in biological and ecological studies is widespread as they can be used to study species distribution, phenology, ecology and behaviour of organisms that are often visually elusive (e.g. marine mammals, bats). This results in extensive data sets (terabytes!) that are scattered in many different locations (e.g. scientific institutes, universities, voluntary researchers). A critical aspect of being able to learn from such large and varied acoustic data sets is providing consistent and transparent access that can enable the integration of various analysis efforts (Roch et al. 2016).

The DASA project (Digital Animal Sound Archive) was established to centralize, preserve, and valorise bioacoustic datasets originating from Belgian research and citizen science initiatives.

DASA started on the 1st of February 2023 and was funded until the 1st of May 2025, was officially extended until the 30th of September 2025 to allow final integration, dissemination, and reporting activities.

3. STATE OF THE ART AND OBJECTIVES

The overarching goal of DASA was to group scattered acoustic recordings into a single infrastructure that safeguards data for future research and provides a reliable reference collection of species-specific sounds. The project also aimed to expand and complement the collections hosted at RBINS, to provide validated datasets suitable for the development of automated identification tools, and to stimulate

ecological studies addressing questions of distribution, migration, and habitat preference. The general DASA database development is fit for purpose for all bio-acoustic data, but within the scope of the project it was set-up for sound recordings of bats, as a proof of concept. In the future, this database structure can be used for sound recordings of other taxa (e.g. marine mammals, insects).

The combination of individual datasets from different providers enhances interpretative capacity by enabling cross-comparison and pattern recognition. This integrative potential was one of the main rationales for developing the archive.

The landscape of bioacoustic data frameworks can be divided into several categories. 1) Bioacoustics metadata models such as Tethys[1] and ecoSound-web[2], 2) Opportunistic/crowdsourcing platforms such as Xeno-Canto[3] and observations.org/waarnemingen.be[4], 3) Citizen science processing pipelines such as BTO acoustic pipeline[5] and VigieChiro[6], and 4) Reference sound catalogues such as ChiroVox[7] and La Sonothèque[8].

Bioacoustics metadata models

Some initiatives focus on metadata models, such as Tethys and ecoSound-web, which provide structured frameworks for describing acoustic data. Tethys is geared towards marine mammals. In addition to a database model, ecoSound-web is also a web application. ecoSound-web is open source software and a running instance of the BioSounds project with additional features. We have analysed the ecoSound-web database model and wanted to express a model that covers the same ground with some extras.

Given the prevalence of simultaneous auto-id/human-id determinations (either standalone or as validation) in ecoacoustics, DASA needed to cover incremental determinations, which is a novel approach. Any agent (humans or agentic determiners such as software or AI) may perform determinations. Determiners have a default fiability ("determination effort") and human determinations are considered to have the highest fiability. In the ecoSound-web database model incremental determinations by humans (interacting through the website interface) are achieved by adding repeat "tags". Novel in ecoSound-web is the delineation of a sound feature: instead of recording only the time at which a bat species starts calling within the record, ecoSound-web allows users to encircle the whole *feature*, a 2D rectangle denoting both the duration (in milliseconds) and the frequency (in KHz). Delineating the precise feature with a drawing tool has not been not deemed essential in DASA as most features (i.e. detections) are extracted by software anyway. Given this human-centrist approach, agentic determinations (i.e. those done by AI or software) cannot be modelled in ecoSound-web.

We also reviewed the ecoSound-web web application. The sound player and feature annotation is very powerful but the interface in general has been found confusing to navigate; perhaps more

[1] https://tethys.sdsu.edu/tethys3/
[2] https://ecosound-web.de/ecosound_web/
[3] https://xeno-canto.org/
[4] https://waarnemingen.be/
[5] https://app.bto.org/acoustic-pipeline/
[6] https://www.vigienature.fr/fr/chauves-souris
[7] https://chirovox.org/
[8] https://sonotheque.mnhn.fr/

features are available after logging in. From this, we identified a functional objective of the DASA platform to be WYSIWYG (what you see is what you get), with plenty of internal links and a powerful faceted search function.

Crowdsourcing platforms

Other frameworks operate via opportunistic or crowdsourced data collection, such as Xeno-Canto and observations.org/waarnemingen.be, where occurrences and recordings are gathered through community contributions. Observations.org is occurrence-centric, Xeno-Canto acoustically-centric. Incremental determinations (i.e. admin validations) are achieved in observations.org by changing the observation species and flagging it as certain.

A key motivation for starting with DASA was the recognition that waarnemingen.be is not designed to handle highly repetitive observations – such as those generated in eco-acoustic monitoring – or to store data volumes reaching tens of terabytes. waarnemingen.be is human-centric and occurrences are added via the interface (ObsIdentify/website). Ultimately this limits the number of species that can be recorded optimally. Due to the biology and behaviour of Chiropterans ecoacoustics is the optimal detection method: a single bat flies around and continuously emits repeat chirps, the downside of which is that we record very frequent repeat detections. Transects with handheld bat detector models lead to fewer recordings than long-term passive acoustic monitoring (PAM) bat detector deployments, but still considerably higher than what typically ends up in “opportunistic” data repositories. Another functional objective of the DASA platform is that it needs to store deployment and project information by default. However, it should be noted that observations.org can be extended with deployment and project metadata in order to capture effort-related information.

Citizen science data processing pipelines

Thirdly, citizen science processing pipelines such as the BTO acoustic pipeline and VigieChiro offer automated analysis of data captured by deployed PAM devices. They can be sent in real or delayed time to these pipelines who return a determination result and/or offer storage in return. An objective of DASA was that the stored sounds should serve as a reference catalogue for such auto-ID pipelines. The architecture of the DASA database and application should be able to achieve this. Other approaches that were considered at the initial stages of the project were : 1) to integrate existing auto-ID software into DASA, 2) to evolve DASA into an acoustic pipeline itself, or 3) to more tightly cooperate with external pipelines.

Reference sound catalogues

Finally, reference sound catalogues such as ChiroVox and La Sonothèque concentrate on curating validated examples of species vocalizations. By collecting the sounds centrally, users should easily find the sound for a specific species.

Based on the above information, the DASA partners decided that that the project developments should aim for a website where Belgian batworkers (both citizen scientists and professionals) can submit bat recordings, together with (if still available) the audio files, gathered by bat detectors (either in transects or via PAM). It should be able to differentiate between auto-ID and human ID and should allow for specialist redeterminations (both citizen scientists and professionals), hence multiplying the effort of the citizen science. The underlying occurrence data should be stored systematically and be

available for further scientific analysis and be feedable into auto-ID software. It should be easy to retrieve representative sounds for a specific species.

The DASA approach aligns with international efforts but places its emphasis on a national-scale integrated database and a web application specifically designed to host Belgian datasets of bat acoustic recordings. While the logo, scope, datasets and interface of the system focus on bats , the system was developed to remain extensible to other taxa and acoustic use cases in the future, while keeping the same underlying database.

Nevertheless, it turned out that DASA contains some determinations on non-bat species, but these are all agentic determinations (mostly by Tadarida and Sonochiro) in datasets made to monitor bats, not other animal groups.

The objectives of the project led to the need of the following functional components. Data provided by partners had to be imported into a unified database. A web application had to be developed to make the data searchable, browsable, and citable, while also offering space for collaborative validation. Storage solutions were to be implemented both on-site and in long-term backup facilities. Quality assurance mechanisms had to guarantee the reliability of metadata and associated recordings. Finally, dissemination of curated datasets was foreseen through integration with RBINS portals and international platforms such as GBIF.

Guiding principles and legislation with relation to data management and infrastructures are the FAIR (Wilkinson et al., 2016; Findable, Accessible, Interoperable and Reusable) principles, the European Open Data Directive and the Interoperable Europe Act. We have designed DASA to align with the FAIR principles by 1) providing persistent, findable and adequate metadata with 2) links to data (wav files) using known protocols; 3) conforming to the GBIF species list vocabulary and providing clear definitions of our internal labels via our API and 4) describing accessibility criteria (in a data policy), study purpose, location, habitat, detection and determination effort in our database model and expose them as lineage information in the metadata. The European Open Data Directive mandates that public sector information should be open by default, even for commercial purposes as much as possible, and by dynamically exposing data in domain specific standards both in bulk form as well as individually (via APIs). DASA is designed to follow these requirements. Finally, the Interoperable Europe Act specifies that IT initiatives of Member States governmental bodies should keep cross-border interoperability in mind, and preferably make use of each other's implementations. In this spirit, open source solutions (both as building blocks as well as final products) are to be preferred. Furthermore, DASA can be translated, re-instanced or skinned to open it to other species groups as well and terminology is kept neutral (e.g. include establishment types not applicable to bats but to other species groups, such as 'introduced and breeding')

4. METHODOLOGY

4.1 Database design and data storage approach

The proposal already contained a preliminary database model. The final database model was designed based on an analysis of existing (database) models (i.e. ecoSound-web, waarnemingen.be and Tethys) and interviews of the data providers, especially RBINS. The model went through a few iterations, finally settling on a model with 54 tables. Tables model specific "entities", who store rows ("instances") of unique information that describe the whole methodology of bat acoustic research. Core entities are Project, Collection, Deployment, Detection, Occurrence and Determination.

Project: The project for which audio files have been collected. This is optional for a collection.

Collection: A group of deployments that should remain together: e.g. monitoring project, volunteer's archive.

Deployment: Making use of an acoustic detector in the field for a certain time, either fixed (tree, windmill, mooring,…) or on a platform (human, animal, car, ship, submarine).

Audio: A digital audio file, either a sound file (wav) or a image rendition of that sound (sonogram). Audio files describe whole deployments or single detections. Detections cover one or more sound pulses. Detections can have either no sound, or one actual file.

DetectionEffort: Device/software settings, protocols, survey goals, calamities or field conditions that influence the reason why a subset of the available data of a deployment has been captured or stored.

Detection: Capturing one or more acoustic signals (pulses) during a deployment of a certain duration, and frequency range by an agent, i.e. a person or by software (on the pc desktop or in the detector). In practice, a detection is never done by a person, always by software.

Occurrence: The presence of an individual (or group of) organism(s) in time and space, of a (the same) specific species, sex or lifestage**.** It is entirely possible a detection found more than occurrence.

Determination: Assigning an occurrence to a source (bat, marine mammal, noise,…) by an agent (person or software). Multiple determinations can be redone over the same occurrence, by different agents (software/persons), keeping an audit trace.

Supportive tables are feature of interest and type, organisation, site and site type, device instance, device model, determination source, fiability type, user account on website etc. We highlight four approaches:

- We distinguish between the feature of interest and the deployment carrier. The feature of interest is the habitat under consideration for the deployment study. Typical example: "Forest clearing in Angleur". The deployment carrier is the thing the detector is attached to. It can be a static site (PAM) in the feature of interest, or it can transect through the feature of interest. Sites are trees, masts,... and moving carriers are persons, cars or boats.
- Determination sources are grouped in animal sources (different animal species, primarily bats, but also crickets and rodents), human sources (piledriving noise, windmills, cars,...), physical

sources (wind, waves, sudden sediment movement,...) and unassigned sources (noise, probably an interesting signal of interest for the study but unattributable; a truly unknown noise).

- Fiability type denotes determination effort and in practice is a categorisation of the determination method. We distinguish the following fiabilities, ordered by ascending quality: classifier (Auto-Id); likely, based on known bias in auto-id software; likely, based on circumstantial detections; human judgement (unspecified); human judgement (sound measurements); human judgement (sonogram); human judgement (auditive); human judgement (combined methods); hand sample; DNA sample. When there is a repeat determination for a single occurrence, with higher fiability quality, it is promoted. Validations done by validators on the website offer an additional level of fiability.
- In agreement with the Follow-up committee, we identified the following data confidentiality levels:
    - Data cannot be reused for any purpose. This protects commercial interests of the data, and can be used when consultancy firms deposit data in DASA. Such data can be seen, but not downloaded in bulk.
    - Data can only be reused for scientific purposes by the DASA group. The DASA group consists of DEMNA (Namur), INBO (Brussels), IGBE-BIM (Brussels), MNHN (Paris, France), Natagora (Namur), Natuurpunt (Mechelen), RBINS (Brussels).
    - Open data, public domain (CC0)
    - Open data, with attribution (CC-BY)
    - Open data, attribution, non-commercial (CC-BY-NC)

  In addition, individual determinations and deployments can be hidden by attaching an embargo period. This is necessary to protect commercial data, in case of deployment at a roosting site or for very rare or sensitive species. All details can be found in https://dasa.naturalsciences.be/content/dasa_datapolicy.pdf.

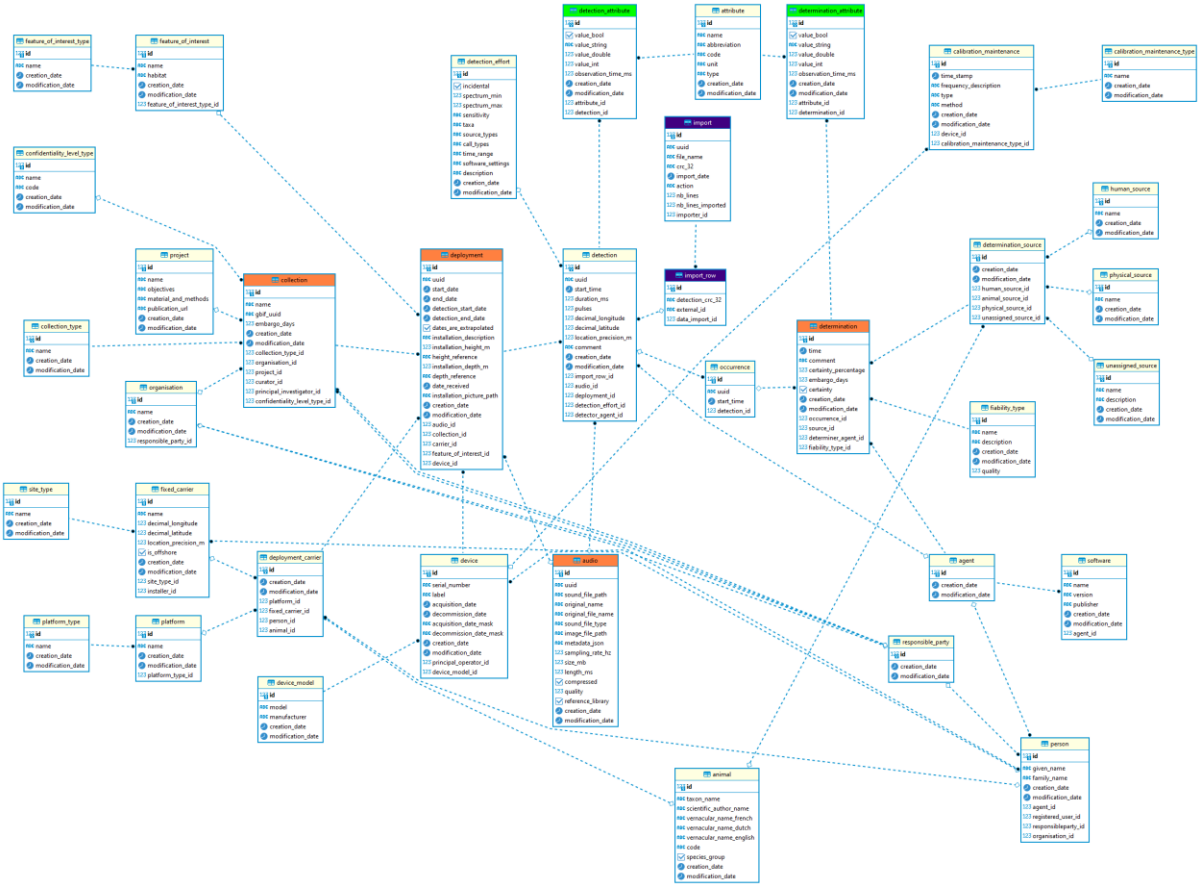

Figure 1. Database model of DASA. Tables with orange headers are core tables, green headers attribute tables whereas purple headers contain information on data imports.

A meeting with the data responsibles of observations /waarnemingen.be at Natuurpunt was organized fall 2023. The topic was the DASA data model and a possible data exchange with observations /waarnemingen.be. A possible avenue for later exchange was identified by sharing data on GBIF. A crucial outcome of this meeting was the addition of the Occurrence concept (instead of just Detection) as the core entities on waarnemingen.be and GBIF are occurrences.

For storage, a dedicated on-premise infrastructure was put in place at RBINS. 20 terabytes were reserved in a central Network Attached file Server (NAS), of which about 10 are currently in use. Of this 20 TB, 12 TB is the main storage with off-premise backups, and 8 TB is used as temporary scratch storage. A 20TB intermediate desktop storage was also acquired to centralise all separate disks provided by the partners (RBINS, Natuurpunt and Natagora). In addition, cold storage was arranged through the BELSPO Long Term Storage facility, which ensures redundancy via a tape system. A cold storage policy was adopted for all files. Very common sounds may be kept out of the primary storage at RBINS but remain retrievable from the LTS for meta-analyses. Exploratory discussions were also held with VigieChiro in France concerning possible long-term backup arrangements and integration of classifier tools.

4.2 Database and import procedure development

Importing wav files happens in a two-step process: first the metadata, then the data. Wav files are submitted together with a completed metadata template, which is an excel sheet documenting

everything needed for the database. This template went through ten iterations between July 2023 and March 2024 in order to capture essential information while remaining practical for providers. A few changes were made as a result of the final follow-up committee meeting and the final workshop in June 2025.

To guarantee data quality, an automated import script was developed in Java. This tool validates the structure of the metadata file, checks the presence of recordings, verifies coordinate accuracy, and maps species codes to standardized taxonomies. The system enforces 21 mandatory fields and 35 consistency checks, among which duplicate entries or submissions, ensuring robustness of the database. All errors are marked in the excel file (in case there are less than 1000 errors), or in a separate text file (in case there are more than 1000 errors). The script will not proceed to import the data into the database as long as an error remains.

This way of working remains practical until +- 250.000 rows. If the submission is larger, the excel file needs to be split in parts. The import script can be used as a command line tool and has been integrated into the website so that users can validate and import their data by themselves.

A second script (in bash) has been written to import wav files from a physical storage disk. First this script indexes all .wav and .raw files on the intermediate desktop storage and writes filename/filepath key/value pairs to a textual index. It then verifies each audio entry in the database whether the actual wav file has been imported before (there is a flag in the database for this); if not, it searches the index for the filename, grabs the file given in the filepath and copies it to the central RBINS NAS storage. Subfolders are created per deployment.

4.3 Data acquisition and validation

Data acquisition was carried out through close collaboration with Natagora, Natuurpunt, and RBINS. Submissions were delivered on hard disks, with accompanying metadata delivered through an Excel-based entry form developed by RBINS. The submissions were received in the course of 2023 and 2024. In some cases, the corresponding sound files could not be recovered; nevertheless in this case, the metadata was retained.

Basic validation is achieved via automated means and has been described in the previous section.

Collaborative validation was designed as a central functionality of the web application. The system distinguishes between members, who are able to suggest species identifications, and validators, who can confirm them. Incremental determinations are stored in the database, allowing for the coexistence of human and machine identifications. A consolidated species list was prepared, covering all Belgian bat species and additional European taxa that may appear due to climate change. These collaborative validation activities were postponed until the web application reached maturity, with plans to combine citizen science modules with external classifiers such as VigieChiro.

4.4 User interface development

The web application itself was built with Spring Boot, React and PostGIS, and provides a password-protected API for programmatic interaction with the data (getting data can be done anonymously). The system is deployed in a Tomcat environment and makes use of a dockerized ElasticSearch instance for searching. It provides overview and detail pages for collections, deployments, and detections. For each detection, the application displays a sonogram, a map, and a download option for the audio file.

Account management, faceted search, a statistics dashboard and data import/export functions are included, and the interface was translated into English, French and Dutch while adopting the RBINS house style.

Dissemination of data was prepared through exposure as GIS data (WMS/WFS), and publication to GBIF. Both are achieved via indexed materialized views.

We assessed the needs of prospective users ("user stories") by sending a Google form survey to the follow-up committee. These recommendations were translated into features and added as tasks in a Kanban board. The code is deposited in an internal RBINS gitlab instance, and discovered issues were noted here. The code is not open source, nonetheless managed with git, hence can easily be made open on request as a new repository in github.

The web application was tested in the following ways:

- Automated junit tests for testing the API, which are run during each deployment step.
- Invitation to test the system sent to the core team (partners) on 20 March, 2025
- Invitation to test the system sent to whole follow-up committee on 2 April, 2025. Bugs and recommendations could be reported in an online spreadsheet.
- The attendants of the last follow-up committee and the final workshop were given the time to explore a test website by themselves. The same spreadsheet to report bugs and recommendations was used.

5. SCIENTIFIC RESULTS AND RECOMMENDATIONS

The DASA project aimed at developing digital infrastructure to safeguard bio-acoustic data and facilitate future scientific research, as such there are no new scientific results that were achieved during the project lifespan. However, the project partners succeeded in building a unified bio-acoustic database supported by a standardized metadata model. The acoustic datasets of the project partners (1.997.170 recordings of bats) were entirely integrated into the database. Around 1 million of these recordings are directly linked to audio files. Some files are matched with metadata through fuzzy timestamp associations. Despite the overall success, data losses were unavoidable, particularly due to hardware failures affecting Natagora submissions. Metadata was nevertheless preserved even when recordings were irretrievable, in order to retain information on detections and deployments.

The approach based on metadata forms proved more reliable than relying on automated GUANO extraction, which was found inconsistent across recording devices. GUANO is a metadata format for bioacoustics, that can be embedded in the wav file. We have created a unified GUANO and basic wav metadata extractor that combines all known tools (tagtools, mediainfo, exiftool, R bioacoustics and emu) in a docker container and picks the best output. The system of validation roles and incremental identifications lays the foundation for collaborative curation once the platform is opened to contributors. Containerized classifiers, including Tadarida and methods described by Roemer et al. (2021), were tested and showed reasonable performance in distinguishing noise from species groups; it does not distinguish separate species. We have integrated the Tadarida software suite in one dockerized pipeline. We may use it in the future to separate biologically relevant noise from environmental (non-biota) sound sources.

Several recommendations emerged from the project. First, future citizen science participation will require a simplified submission workflow, ideally integrated in the web interface, rather than relying solely on metadata forms. Contacts with the Walloon volunteer network is promising. Manual metadata remains important, yet is error-prone, nevertheless full reliance on embedded metadata (GUANO) and GPS coordinates should be approached carefully as not all models include these (reliably). Second, even common species and noise data should never be discarded completely, as their potential value for meta-analyses and ecological modelling is considerable. Third, there is still a need for better reference data: visual confirmations of recorded sounds would fulfil the reference collection objective completely.

Finally, a continued collaboration with external initiatives such as VigieChiro is essential to maintain classifier validation and long-term sustainability of the archive. Integration of existing auto-ID software into DASA, or letting DASA evolve into an acoustic pipeline itself have been identified by the Follow-up committee and the data partners as undesirable. Cooperation with external pipelines entails interoperability of their standards and hence better coverage of Belgium in the European ecoacoustics landscape.

RBINS already has a contract with MNHN for data sharing, and full backup in VigieChiro (in fact, the IRODS system of the IN2P3 in France) is included in the follow-up project's proposal. DASA (and the wider need for bat monitoring data) has become a cornerstone of the RBINS data storage acquisition strategy. We estimate there to be around 50TB of volunteer data and more than 200TB of professional data in Belgium.

6. DISSEMINATION AND VALORISATION

6.1 Workshops

Dissemination activities took several forms. Several workshops were organised to reach out to future users of the DASA platform, aiming both at professionals and amateur bat researchers. A first workshop was organised on June 25, 2025 at the Institute of Natural Sciences. This was targeted at professional users and the project partners. The aim was to present the developments and have a 'hands-on' training to get acquainted with the data importing procedure. 15 researchers from research institutes, consultancy agencies, and partner institutes joined the workshop.

To introduce the platform to volunteers and citizen scientists, we organised workshops at the project partners Natagora and Natuurpunt. With their respective working groups Plecotus and 'de vleermuizenwerkgroep', they have an elaborate network of amateur bat researchers. These workshops were designed to provide training on account creation, use of the metadata form, and preliminary validation exercises with small test datasets. A first workshop of this kind was organised in Namur, in collaboration with Natagora, on September 27, 2025. A workshop at Natuurpunt in Mechelen will be organised later this year.

6.2 Presentations

The project results were presented at several conferences and symposia:

- Oral presentation by Bob Vandendriessche during the Belgian Bat Symposium in Halle, 7 October 2023
- Poster presentation at the International Conference on Marine Data and Information Systems (IMDIS), 27-29 May 2024 in Bergen (Norway)
- Oral presentation at the Open Science Café: Passive Acoustic Monitoring at the Institute of Nature and Forestry Research, on 19 November 2024
- A dedicated talk at MNHN (Muséum national d'Histoire naturelle) in Paris on 10 March 2025
- Casual science outreach at the Citizen Science Fair in the Museum grounds at RBINS on 18 June 2025
- DASA was introduced as a data repository to Natagora volunteers during an acoustic training of 6-7 September 2025
- Bat acoustics workshop organised by Natagora on 27 September 2025.

6.3 Press releases

The DASA website will be officially launched in November 2025 on https://dasa.naturalsciences.be. To draw some extra attention to this, a press release will be sent out by the project partners.

6.4 Follow-up projects

During the DASA project, we aimed to meet the storage and data management needs for both bio-acoustic professionals and volunteers. To valorise these developments and expertise, we prepared a follow-up project for the BELSPO Policy4Science call. The proposal, entitled OSIRIS (Organized Storage of Images and Recordings for Interoperable Science), was submitted in May 2025. It aims to extend the scope of DASA by addressing broader research data management needs, including marine acoustics, noise monitoring, ROV images and videos, and additional bioacoustic domains. The continuity of DASA itself is foreseen in the form of ongoing data acquisition, especially from volunteers, and maintenance of the website.

6.5 Follow-up committee

The follow-up committee consisted of researchers from NGOs and governmental bodies, ecological consultants, and citizen scientists. The committee members received an invite during the proposal stage, two more were contacted after the project started, and one member joined later on after hearing about the project. The project partners met with the follow-up committee during several occasions during the course of the project. During the meetings, the project partners presented the progress on the project's deliverables and the web application, and organised topical discussions. Topics covered were: 1. What is the overall goal of the project, hence which website functionalities are most needed from an end-user perspective? 2. What levels of confidentiality are needed? 3. What constitutes noise and how should it be managed (considering future prospects on sound analysis)? 4. What metadata should be mandatory and how to cater to professionals/volunteers?

During the last two meetings, attendants had the time to explore a test website and provide feedback. This was organised as a user acceptance test session. We launched two Google form surveys. One to identify the website's features by surveying the user needs, the other to ask how much data persons have. Two Google spreadsheets were used; one to assess the user landscape and stakeholders, the

other to report bugs and recommendations when testing the DASA web application. The stakeholder spreadsheet has been circulated many times.

The follow-up committee meetings were held on:

- October 6, 2023 (kick-off meeting)
- April 25, 2024
- December 11, 2024
- April 24, 2025
- the follow-up committee also participated to the workshop on June 25, 2025

6.6 Other valorisation efforts

Collaboration with external stakeholders was pursued throughout the project. Contacts with INBO focused on possible alignment with the Camtrap DP Extension[9] developed under TDWG. The DASA data model has been studied by the Extension's authors. Close collaboration with Natuurpunt and Natagora was maintained, both for the DASA submissions and for the preparation of follow-up projects.

RBINS data was submitted to GBIF in this dataset: https://www.gbif.org/dataset/ac4f02e9-ce6a-4132-892a-b478bc79ed14. The data is organised according to best practices for organising camera trap data[10], a procedure that can be largely applied to acoustic data as well. Data from the partners Natagora and Natuurpunt may be disseminated as well on GBIF; the procedure is very similar and takes little time to implement.

7. PUBLICATIONS

Vandenberghe T., Brabant R., Laurent Y., Brabant C., Vandendriessche B., Willems W., Degraer S. & Lagring R., 2024. Digital Animal Sound Archive: a collaborative repository for bio-acoustics In: International Conference on Marine Data and Information Systems - Proceedings Volume, vol. 80, pp. 165-167. Miscellanea INGV.

Laurent Y., Vandenberghe T. (2025). RBINS bat sound collection. Royal Belgian Institute of Natural Sciences. Occurrence dataset https://doi.org/10.15468/jgudzj accessed via GBIF.org on 2025-10-20.

[9] https://wildlabs.net/discussion/safe-and-sound-standard-bioacoustic-data
[10] https://docs.gbif.org/camera-trap-guide/en/

8. ACKNOWLEDGEMENTS

The authors wish to thank the Belgian Federal Science Policy office (BELSPO) for the funding of the project under the BRAIN-be 2.0 programme. We value the input of everyone throughout the team and follow-up committee meetings.

DASA partners + affiliations:

Name	Surname	Organisation
Brabant	Claire	Natagora
Brabant	Robin	RBINS
Laurent	Yves	RBINS
Vandenberghe	Thomas	RBINS
Willems	Wout	Natuurpunt

Follow-up committee members+affiliations:

Name	Surname	Country	organisation	sector
Adriaens	Véronique	Belgium	Sertius	consultancy
Casier	Louis	Belgium	CSD Ingenieurs	consultancy
Conings	Bram	Belgium	VLIZ	research
Corens	Peter	Belgium	Tractebel	consultancy
De Bruyn	Luc	Belgium	University of Antwerp	University
Gyselings	Ralf	Belgium	Research Institute for Nature and Forest (INBO)	research institute
Hermans	Claire	Belgium	Wageningen University	University
Lefevre	Alex	Belgium	Vleermuizenwerkgroep Natuurpunt	environmental NGO
Newson	Stuart	UK	British Trust for Ornithology, Data Science and Bioacoustics	research
Opstaele	Bart	Belgium	GreenSpot	consultancy
Roemer	Charlotte	France	Muséum national d'Histoire naturelle	research
Smits	Quentin	Belgium	DEMNA	governmental agency
Vandendriessche	Bob	Belgium	Regionaal Landschap Houtland & Polders vzw	Environmental NGO
Van De Sijpe	Marc	Belgium	Vleermuizenwerkgroep Natuurpunt	environmental NGO
Van der Wijden	Ben	Belgium	BIM	governmental agency
Verkem	Sven	Belgium	N8	consultancy

9. REFERENCES

Berwick, R. C., Okanoya, K., Beckers, G. J., & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. Trends in cognitive sciences, 15(3), 113-121.

Brabant, R., Laurent, Y., Poerink, B. J., & Degraer, S. (2019). Activity and behaviour of Nathusius' pipistrelle Pipistrellus nathusii at low and high altitude in a North Sea offshore wind farm. Acta Chiropterologica, 21(2), 341-348.

Hawkins, A.D. (1986). Underwater sound and fish behaviour. In The behaviour of teleost fishes (pp. 114-151). Springer, Boston, MA.

Richardson, W. J., Greene Jr, C. R., Malme, C. I., & Thomson, D. H. (2013). Marine mammals and noise. Academic press.

Roch, M. A., Batchelor, H., Baumann-Pickering, S., Berchok, C. L., Cholewiak, D., Fujioka, E., ... & Van Parijs, S. (2016). Management of acoustic metadata for bioacoustics. Ecological informatics, 31, 122-136.

Roemer, C., Julien, J.-F., Ahoudji, P. P., Chassot, J.-M., Genta, M., Colombo, R., Botto, G., Negreira, C. A., Djossa, B. A., Ing, R. K., Hassanin, A., Rufray, V., Uriot, Q., Vigie-Chiro Participants, & Bas, Y. (2021). An automatic classifier of bat sonotypes around the world. *Methods in Ecology and Evolution*, 12(2), 2432–2444. https://doi.org/10.1111/2041-210X.13721

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18