Towards standardized soundscape recordings and analyses in Earth System Sciences

Kevin FA Darras[1], Frederik Pitz[1], Dánnell Quesada-Chacón[1], Felix Zichner[1], Paula Phan[1], Jan O Engler[1], Anna F Cord[1]

[1]Chair of Computational Landscape Ecology, Helmholtzstrasse 10, TU Dresden, 01069 Dresden, Germany

Date of submission: 31 March 2023

This work has been funded by the German Research Foundation (NFDI4Earth, DFG project no. 460036893, https://www.nfdi4earth.de/).

Pilot title	PAMbase: a Repository of Soundscape Recordings to study Earth's Phonosphere
Project Duration	12 months
Contributors[1]	Kevin FA Darras: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing Dánnell Quesada-Chacón: Investigation, Software, Writing – Original Draft Preparation Felix Zichner: Investigation, Methodology, Validation, Writing – Original Draft Preparation Frederik Pitz: Investigation, Validation Jan O Engler: Conceptualization, Funding Acquisition, Methodology, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing Anna F Cord: Conceptualization, Funding Acquisition, Supervision, Writing – Review & Editing
DOI	10.5281/zenodo.7785327
Corresponding author	Kevin FA Darras, Anna F Cord

Abstract

Passive acoustic monitoring is an effective method for tracking a variety of phenomena in the Earth System Sciences, whether they are devastating earthquakes (lithosphere), singing whales (biosphere) or human transportation (anthroposphere). However, we lack common databases and software tools for an integrated analysis of the acoustic information gathered across the different spheres. We aimed to deliver an operable software tool for inter-disciplinary acoustic

[1] Based on the CRediT Contributor Roles Taxonomy: https://credit.niso.org/

analyses, and to engage the community around a roadmap for standardized workflows in passive acoustic monitoring. We present the front- and back-end of our tool based on an open-source ecoacoustics platform along with the functionality to allow the participation of the general public and the execution of deep learning algorithms to automatically identify bird species in sound recordings. We also 1) surveyed user needs across the NFDI4Earth and NFDI4Biodiversity consortia; 2) present guidelines for data quality control, standardisation, and sharing; and 3) propose future directions for the technological and methodological development of common databases and software tools. Although we were not able to comprehensively assess or federate the Earth System Science community, we present clear improvements for the software-based acoustic data workflow used by the biogeographical research community.

I. Introduction

Sound is a vibration traveling through liquid, solid, or gaseous media, created by a variety of moving sources across a wide range of frequencies spanning infrasound (between 3-10 mHz and 20 Hz), human-audible sound (between 20 Hz and 20 kHz), and ultrasound (above 20 kHz). Sounds can be classified according to their source into geophony (geophysical processes), biophony (living organisms), and anthropophony (humans or their activities). It follows that we can infer the identity, location, or state of sound sources over a wide range of applications across disciplines in the Earth System Sciences. Here, we follow integrative attempts to categorise Earth system spheres into the lithosphere (crust and uppermost solid mantle), the pedosphere (soil), the atmosphere (gas layers enveloping the Earth), the cryosphere (all solid water), the hydrosphere (all liquid water), the biosphere (habitats with living organisms), and the anthroposphere (human-made habitat) (Steffen et al., 2020).

Analogously to the spheres concept, and to foster cross-sphere integration, we introduce the term “phonosphere”, which comprises the passively sensed acoustic domain transcending all the spheres, and describe its state. Low-frequency acoustic waves are commonly recorded using seismographs, geophones, microbarometers, or accelerometers. Vibrations in the lithosphere inform us about its structure (Rost and Thomas, 2002), indicate failure events in hydrocarbons exploration (Eisner et al., 2012), and can be used to predict earthquakes (Nikolaev et al., 2005, p. 20). The lithosphere is coupled to the atmosphere such that infrasound propagating in the atmosphere can also be used to monitor not only earthquakes, but also volcanoes, the atmosphere itself, and the oceans, much in the spirit of studying the Earth as interconnected systems (Hedlin et al., 2012). Also the hydrosphere is thoroughly investigated acoustically, and by virtue of the high acoustic transmission speed and range in water, long-range surveillance of human and animal activities and meteorological and geological events such as storms and volcanoes are possible (Duarte et al., 2021). Even frozen water in the cryosphere is acoustically monitored to track rapid changes due to global warming, spawning new approaches for monitoring thawing permafrost (Cheng et al., 2022) and tremors in glaciers (Liashchuk et al., 2021). Higher-frequency audible or ultrasonic frequencies are recorded with microphones to monitor human activities and the effects of their sounds on human health and well-being (Mohamed et al., 2021) as well as wildlife (Blickley and Patricelli, 2010).

Microphones are also commonly used to monitor the state and dynamics of the biosphere - mostly for animal activities and occurrences and ecosystem health (Gibb et al., 2018). Note that active acoustic sensing methods, which use acoustic emitters to probe the Earth system spheres, are covered elsewhere (Benoit-Bird and Lawson, 2016).

Passive acoustic monitoring sensors transduce mechanical energy (from the acoustic waves) to electrical energy which can be recorded for visualization, analysis and interpretation. This common underlying principle calls for the use of shared, sound-specific tools such as archival databases, data validation, standardisation, and dissemination protocols, as well as visualisation, annotation, and analysis techniques. Potentially, such a unified methodological approach can transform the field of passive acoustic monitoring in Earth System Sciences by enabling transdisciplinary syntheses across spheres and stimulating systems thinking. From a technical point of view, this unification would be achieved on the information level, with a database that collects and organises meta-data and primary data from research projects, but also on a software level, with a web-based tool that enables collaborative work on and sharing of acoustic datasets to generate secondary data and insights using dedicated functions. Achieving the construction of such a collaborative web-based space and database is the aim of the PAMbase project.

At the time of writing, we are aware of collaborative projects in atmospheric infrasonic monitoring based on the International Monitoring System, that produce secondary data available for a wider range of users (Hupe et al., 2022, p. 200). Worldwide acoustic monitoring research projects also exist for monitoring urban settings, such as the “Urban Soundscapes” project (De Coensel et al., 2017), and a range of software tools and collaborative platforms also exist for federating ecoacoustic research in the biosphere (Darras et al., 2020b). Arguably, the project that comes closest to the aims of PAMbase is the Earthsound project (“The Earthsound Project,” n.d.), which achieves accelerated streaming of seismic and infrasonic data from multiple locations on the globe. However, the analysis features are limited, and most streams are managed by a single person. Finally, no project explicitly adheres to the FAIR principles (Wilkinson et al., 2016), making integrative, collaborative work challenging as of now.

The original aim of this roadmap is to systematically identify the needs of users of passive acoustic monitoring data to guide the development of software-based collaborative tools, and to propose guidelines for managing data quality, standardisation, and sharing issues. The original aims have not only been reached in theory, but were also implemented in practice with the adoption and extension of an open-source ecoacoustics platform (Darras et al., 2020b). We report our results on the progress of developing Earth System Science-specific features on that platform, discuss research challenges and gaps, outline the relevance of our work for the community and NFDI4Earth, and close this roadmap with an outlook of future needed technological advances for the development of a passive acoustic monitoring database for Earth System Science.

II. Results

a) Implemented Solution

The technical implementation is guaranteed by the ecoSound-web platform (Darras et al., 2023), an open-source internet platform originally developed externally for collaborative ecoacoustics projects, and currently co-developed with PAMbase resources and members.

For deliverable 1 (“Technical Operability”) and 2 (“Community Engagement and Roadmap”), all goals were reached. Most aspects of deliverable 2’s “ii” goal, which is to develop “guidelines for quality control, data standards and data sharing”, were aditionally technically implemented on ecoSound-web.

WP1: Operational front- and back-end

Proposal goal: the open beta version of PAMbase will consist of an operational front- and back-end allowing to set up user and project profiles and upload soundscape files with standardized metadata

ecoSound-web’s **front-end** enables upload, archival, visualization, navigation, filtering, playback, manual annotation, collaborative review, and automated analysis of sound recordings with a focus on soundscapes (Darras et al., 2023). Recordings are organised within searchable collections, which are part of projects (and collections are searchable themselves), for which individual user privileges can be set up. The access of users to specific projects is managed by administrators and project managers (i.e., users with management privileges). Acoustic data are embedded within their spatio-temporal context, through their geographic sites and recording date-times. Locations are displayed on tile-based, interactive Leaflet maps, and recordings can be displayed on timelines (Figure 1).

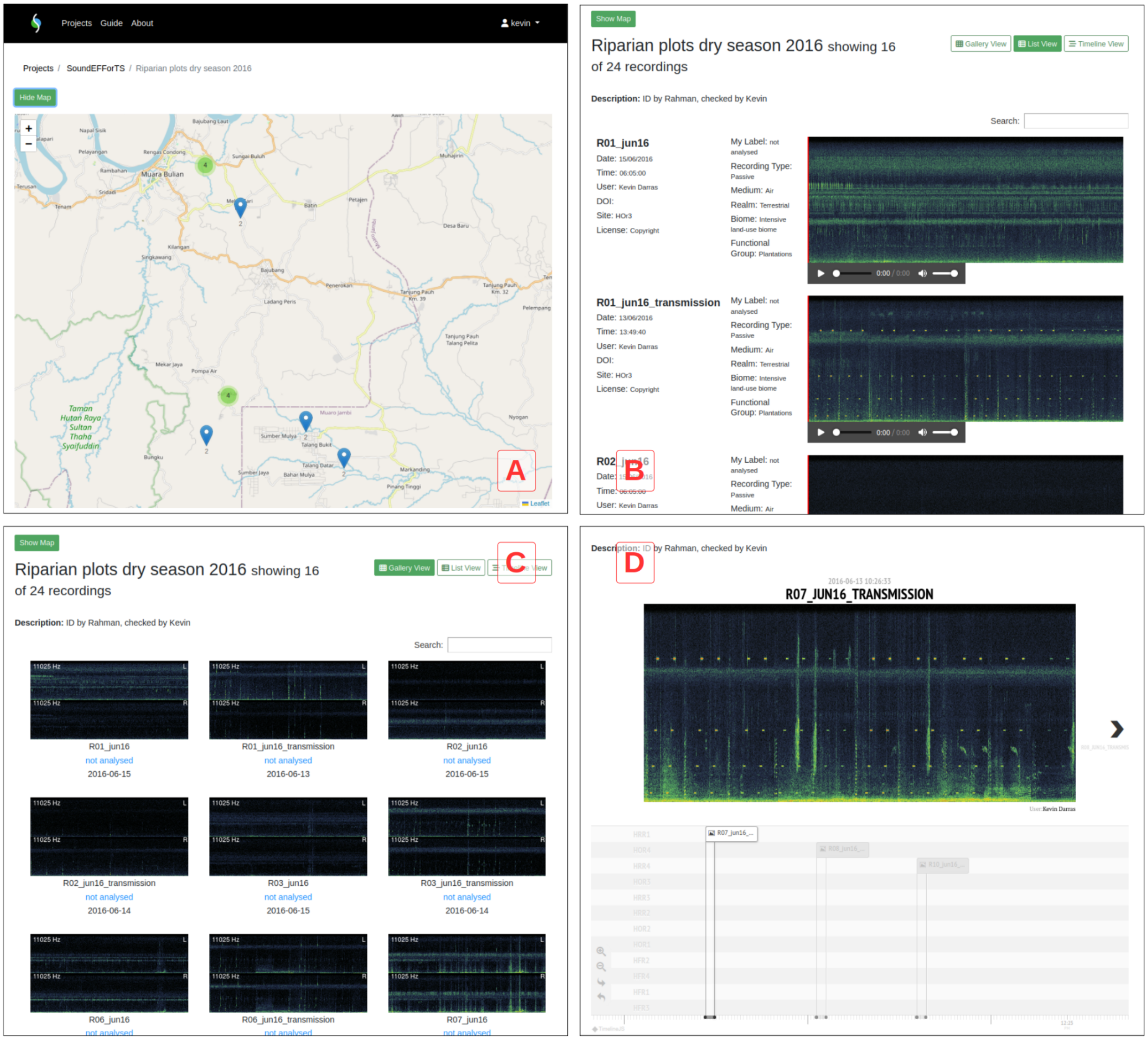

Figure 1: The gallery (C), list (B), and timeline (D) views for recording collections in ecoSound-web, along with the interactive site maps (A). Figure extracted from (Darras et al., 2023).

ecoSound-web's **back-end** consists of code written in PHP 7, Python 2.7 and 3.10, Javascript, JQuery 3.4, Twig 2, CSS and HTML 5. We use a mySQL relational database to structure all data (Figure 2). Sounds and images are processed with Web Audio API, Sox 14.4, Lame, ImageMagick and Scikit-maad 1.3.12. We use a RabbitMQ queue for file processing, Plupload 1.5 as a visual file upload tool, GADM as administrative regions for the sites. JQuery UI 1.12, JCrop 0.9, Bootstrap 4.3, Leaflet, Timeline.js, Bootstrap-selected, Jquery.cookie, DataTables and the Symfony 4 process components are used for managing the scripts execution. Further Python libraries used are: Numpy, Pillow, Audiolab 0.8, Matplotlib, SciPy and Scikit-image. The project is containerized using Docker, which spares software developers the time for installing libraries, the database, and configuring the server. This setup allows developers to run the project on their machines quickly and free of typical installation issues like library version incompatibilities.

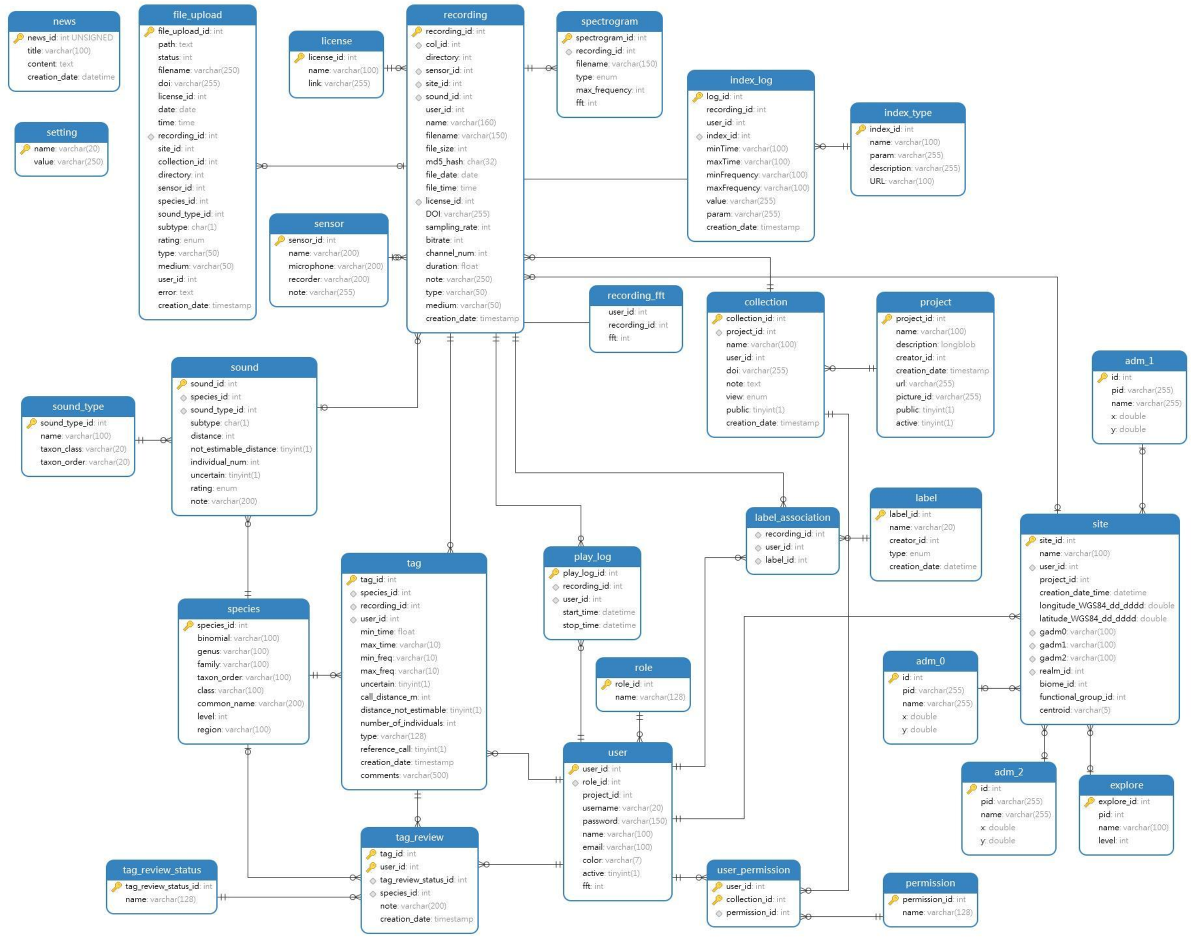

Figure 2: *Database structure with tables, relationships, primary keys, and data types*

WP2: Crowdsourcing

Proposal goal: [module] allowing the participation of the general public

Several access management features enable the participation of the general public, as well as citizen scientists to recording collections and projects, and thus fulfill the requirements of WP2. The access to collections is regulated on two levels (Figure 3): First, collections can be either publicly-accessible or closed. Recordings within public collections can be opened, navigated, played back, and their annotations optionally displayed, by any visitor without requiring login. However, advanced features that would insert new data into the database, such as manual annotation, deep learning models, or acoustic index computation are not enabled for visitors. Second, privileges for closed collections can be regulated for individual users. Recordings within accessible collections can always be annotated by registered users. Annotations of other users can be seen or reviewed by designated users.

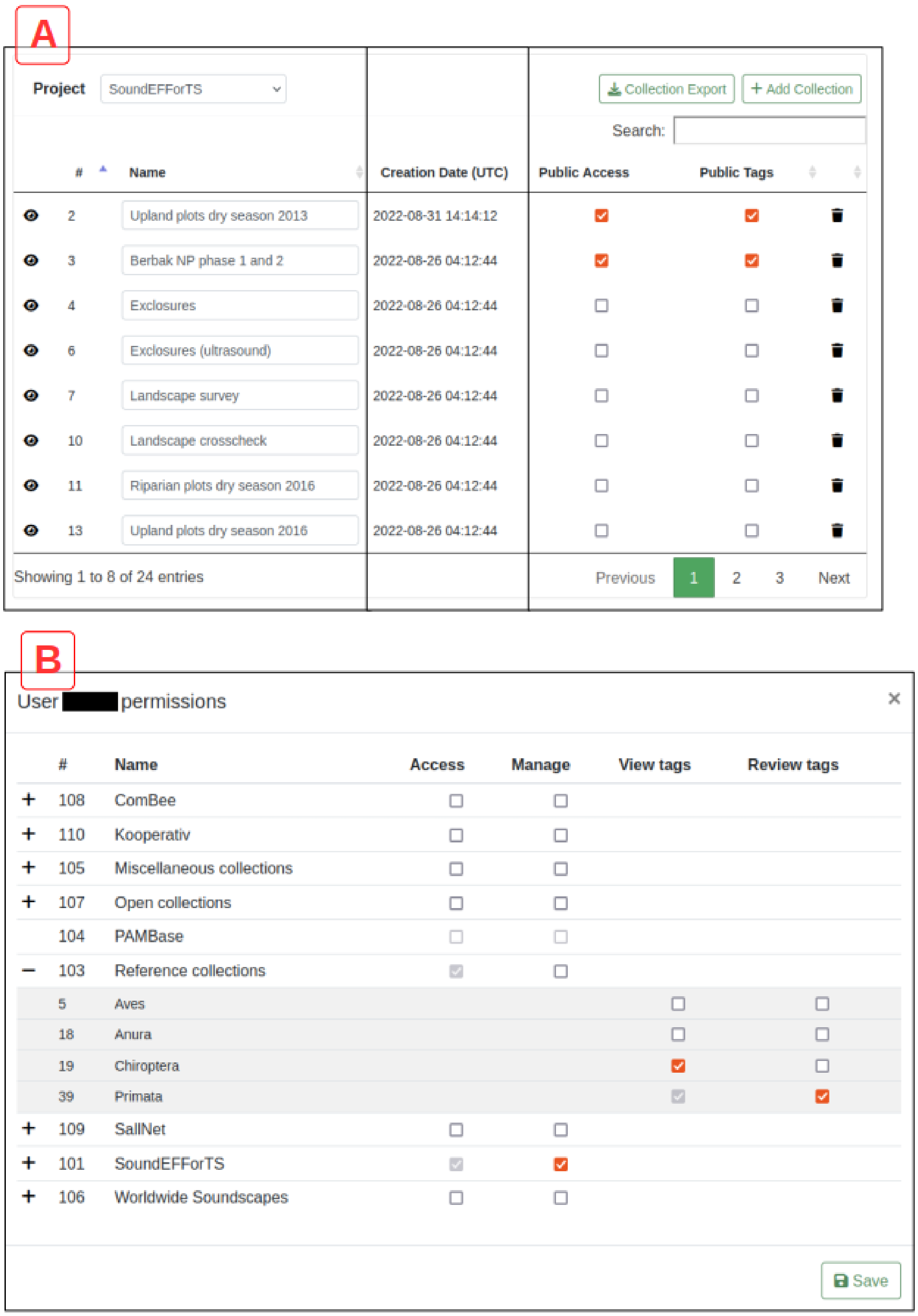

***Figure 3**: Admin interface for (A) managing closed and public collections, with or without public annotations (several columns hidden); (B) managing user privileges for individual collections grouped within projects.*

WP3: Machine learning

Proposal goal: [module] enabling sophisticated data analysis through machine learning techniques

We chose the Tensorflow framework to enable the machine-learning-based, automated analysis of sound recordings. We selected TensorFlow rather than coding-intensive Torch-based approaches or deep-learning incompatible scikit-learn approaches to implement established models for identifying bird call species. The current state-of-the art is BirdNET Analyzer, a deep artificial neural network from the ResNet family (Kahl et al., 2021), which identifies bird calls within 3-second long audio snippets. We can execute BirdNET Analyzer from the spectrogram player (Figure 4) or the admin interface, and the output of the classification is used to generate new annotations. The sensitivity, threshold confidence score, shifting window overlap, and minimum species occurrence frequency threshold can be set or left at their default values. Other execution parameters, such as the geographic coordinates or the calendar week, are automatically extracted from the database to narrow the model output to the range of species that can occur at the specific time and location.

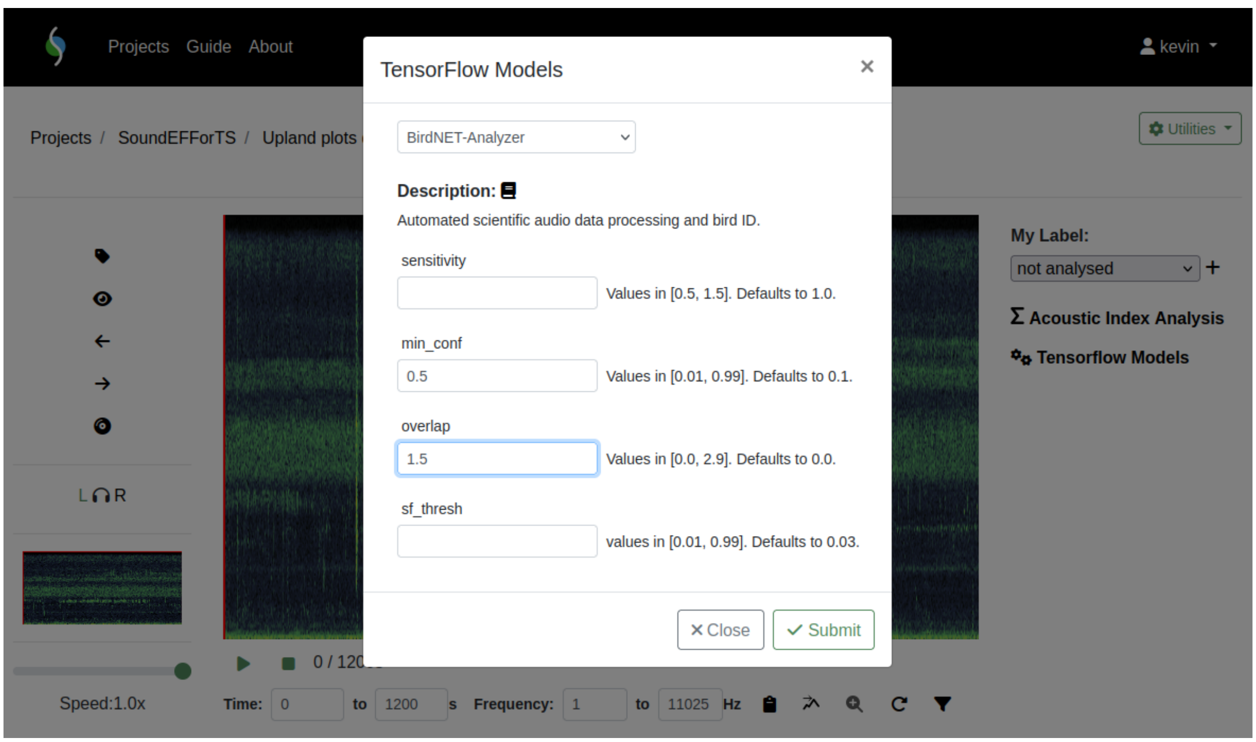

***Figure 4**: Execution options for the TensorFlow-based BirdNET Analyzer. Model output is inserted as new annotations into the recording after execution.*

b) Data and Software availability

- A running instance of ecoSound-web is available at https://ecosound-web.de/.
- The platform is documented in F1000Research: https://f1000research.com/articles/9-1224/v2 under a Creative Commons Attribution License (CC BY 4.0) and registered with the DOI *10.12688/f1000research.26369.2*
- The source code at the time of the F1000Research article publication was archived in Zenodo in the “Second release of ecoSound-web”: https://zenodo.org/record/7603400#.ZBssFI7MJ3E under an Open license, and registered with the DOI *10.5281/zenodo.7603399*
- The latest source code is available in our public GitHub repository: https://github.com/ecomontec/ecoSound-web under the GNU General Public License (v3). We compiled a GitHub Wiki for end users, and development plans are visible in the open projects.

Most of the code contributions to the ecoSound-web repository (which was forked from BioSounds: https://github.com/nperezg/biosounds) were made after adopting the platform for PAMbase (Figure 5).

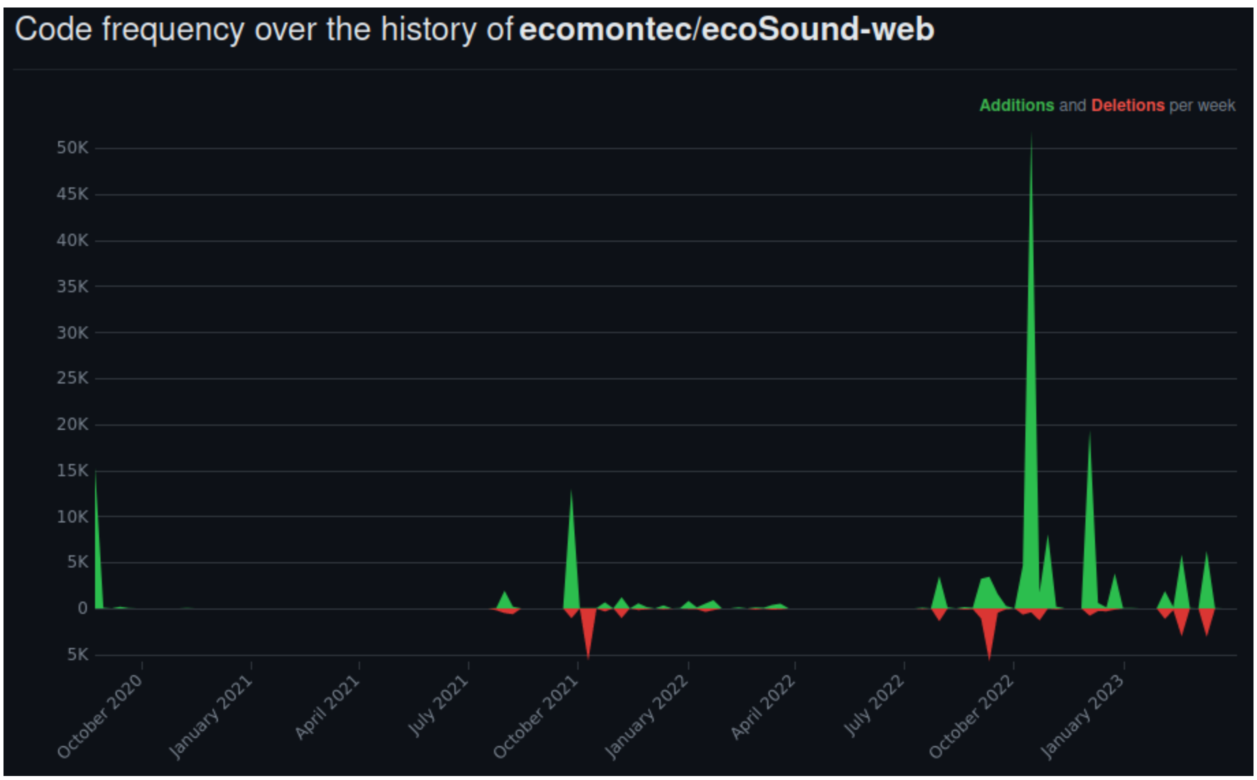

***Figure 5**: Code frequency over the history of the ecoSound-web repository. The time period from the adoption of the platform unti the end of the pilot is outlined in blue.*

c) Innovation and FAIRness

Innovation

ecoSound-web is the most holistic platform for web-based acoustic analyses of passively sampled sound (Darras et al., 2023). First, the platform is not restricted to particular countries or regions (c.f. WildTrax[2]), nor restricted to particular taxa (c.f. Raven Pro[3]), nor restricted to particular biomes (c.f. OPUS (Thomisch et al., 2021)). Second, the platform integrates performant, specialised, and established components by including scikit-maad for computing acoustic indices (Ulloa et al., 2021), BirdNET Analyzer for automatic ID of birds within soundscapes (Kahl et al., 2021), and by linking bird annotations to species searches on XenoCanto (Xeno-canto Foundation, 2012). Third, it offers not only a workspace for managing and analysing soundscapes, but addresses the entire ecoacoustic workflow (Darras et al., 2023) and additionally provides access to species call reference collections such as primates, bats, and anurans, which are traditionally mostly separated (c.f. XenoCanto[4], ChiroVox (Görföl et al., 2022)). Compared to the existing Earthsound project, ecoSound-web offers projects and users management, analysis of audio recordings with AI tools and acoustic indices, annotation management and peer-review, interactive spectrograms, as well as open-source code and published documentation. On the Earthsound project, infrasonic streams are currently down so that it is not possible to check more features (only seismic streams are available). However, advantages of Earthsound include live streaming and a higher level of detail for the sensor metadata.

FAIRness

FAIR principles were explicitly considered with the goal to propose ii) “guidelines for quality control, data standards and data sharing” and the goal to outline (iii) “a concrete time plan for further technological and methodological development during the NFDI4Earth consolidation and advancing phase”.
The “**F**indable” and “**A**ccessible” principles are covered under the Sharing section, the “**I**nteroperable” principle under the Future Directions section, and the “**R**eproducible” principle pertains to the Quality control and Standardisation sections below.

FAIR principles implementation can be assessed on different levels, so we define the following types of data:

- meta-data: information that describes uni-dimensional properties of the acoustic recordings and, by extension, their collections and projects.
- raw data: the acoustic recordings themselves
- secondary data: information derived from the analysis of the raw data, such as annotations, spectrograms, acoustic indices, etc.

[2] https://www.wildtrax.ca/home
[3] https://ravensoundsoftware.com/software/raven-pro/
[4] https://xeno-canto.org/

Quality control

Here, quality control refers to manual processes by which users review the data.

We used a mySQL relational database structure to control the quality of the **meta-data** (Figure 2). However, note that meta-data are descriptive and as such, do not require quality control *per se* beyond coherence checks and standards that are implemented in an automated way with the relational database structure. By defining the data types of the table fields, as well as entry input rules to allow certain ranges of values, sometimes in dependence on values from other fields, we prevent that invalid data are entered. By designating primary keys and relationships between tables, we also prevent the entry of duplicates and make sure by storing data on different levels that no redundant data are saved.

Raw data quality is primarily determined by the signal-to-noise ratio, a pivotal quantitative measure of how cleanly acoustic signals are captured, which determines secondary environmental data (Darras et al., 2020a). Experts usually assess the perceived quality by listening to the recordings and inspecting their spectrograms (in itself a type of secondary data). Accordingly, we propose to assess acoustic recording quality based on the overall signal-to-noise in the recording.

We propose the following classification of recording quality:

1. *technically successful*: no issues with the technical quality of the recording. Signal-to-noise ratio as high as can be expected from the specifications of the microphone.
2. *broadband noise*: intermittent or continuous bands of noise across the entire frequency spectrum, pointing to malfunction of microphone or amplification circuitry. Signal-to-noise ratio usually too low to distinguish signals of interest inside the noisy parts.
3. *silent*: usually continuously silent recording, indicating temporarily water-logged or permanently, mechanically damaged microphone diaphragms. Signal-to-noise ratio very low so that signals of interest can only be discerned with great difficulty or amplification.

The quality of **secondary data**, primarily consisting of annotations, can be checked by peers. Annotations consist of any information that a human wants to affix to a specific spectro-temporal recording region. Often, the sound source type is annotated after manual determination by experts, but automated tools are also increasingly used (Priyadarshani et al., 2018). However, passive acoustic monitoring does not provide any additional (e.g., visual) information about the identity of the sound source (bat, earthquake, car honk) and as such, the identification of recordings by human experts is prone to interpretation errors. Quality checks for annotations are required and implemented with a peer-review system for assessing the identification validity of the sound source by users with equal or superior expertise. Such a validation capitalises on the potential of acoustic data for verifiability (Darras et al., 2019), thereby increasing their credibility. In the future, we will implement productivity- and visibility-enhancing functions to improve the reviewing experience (review archival, visual enhancements for distinguishing review types, etc.)[5].

Note that further secondary data types such as frequency spectra, spectrograms, or acoustic indices are mathematically-defined and do not require quality control if correctly implemented on a technical level.

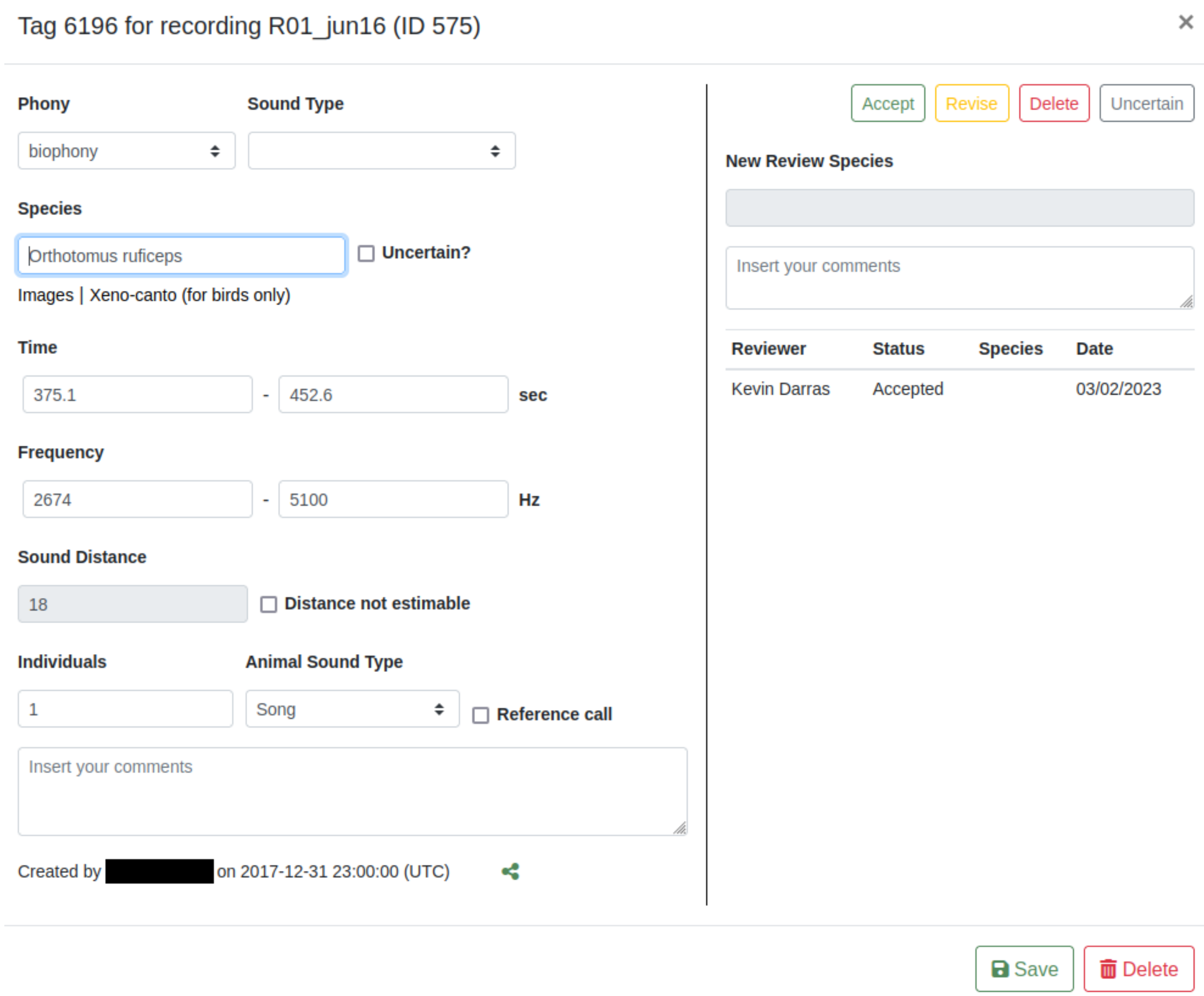

***Figure 6**: ecoSound-web annotation window with right-hand side review pane. Extracted from (Darras et al., 2023)*

Standardisation

Meta-data can be standardised in a variety of ways. The ecoSound-web database is a relational MySQL database (Figure 1) which connects data tables containing different established, standardised data sources stored in fields of different data types.

[5] https://github.com/ecomontec/ecoSound-web/issues/71
https://github.com/ecomontec/ecoSound-web/issues/69

Sampling locations must be unambiguously described and topography information is needed as sound can be sampled at different depths for a given latitude and longitude. Accordingly, we implemented numerical UTM coordinates in decimal degrees format with a WGS84 projection. An additional field for the metric topography (elevation above or depth below sea level) describes the vertical level of the ground or water layer where the sampling happens. Freshwater bodies, whose surface can be situated at positive topography values, additionally require depth information to describe where the sensors were placed relative to the water body surface. An additional field that describes with more precision how far above the ground the acoustic sensors are situated will be useful (not implemented yet) to take sound wave reflection into account, which affects the sound detection space (Darras et al., 2016). We require local times to facilitate the input of recording times, as they are used by the vast majority of users. Standardised UTC times can be automatically derived from the local times and geographic coordinates.

We propose following the IUCN Global Ecosystems typology, comprehensively covering both aquatic and terrestrial systems, to describe in which ecosystem type the recordings were made (Keith et al., 2022) and adhering to the Global Administrative Database (gadm.org/) standard to describe in which administrative region the recordings were made, on the national level (GADM0) and lower levels (level 1 and 2). We implemented IUCN GET ecosystem types and GADM regions at the level of the sampling site (Figure 7). GADM areas also have defined geographic boundaries, so that sites without coordinates can be displayed at the centroid of their administrative area. The integration into the Earth system sciences framework also required the inclusion of a field identifying the spheres of collections (Figure 7).

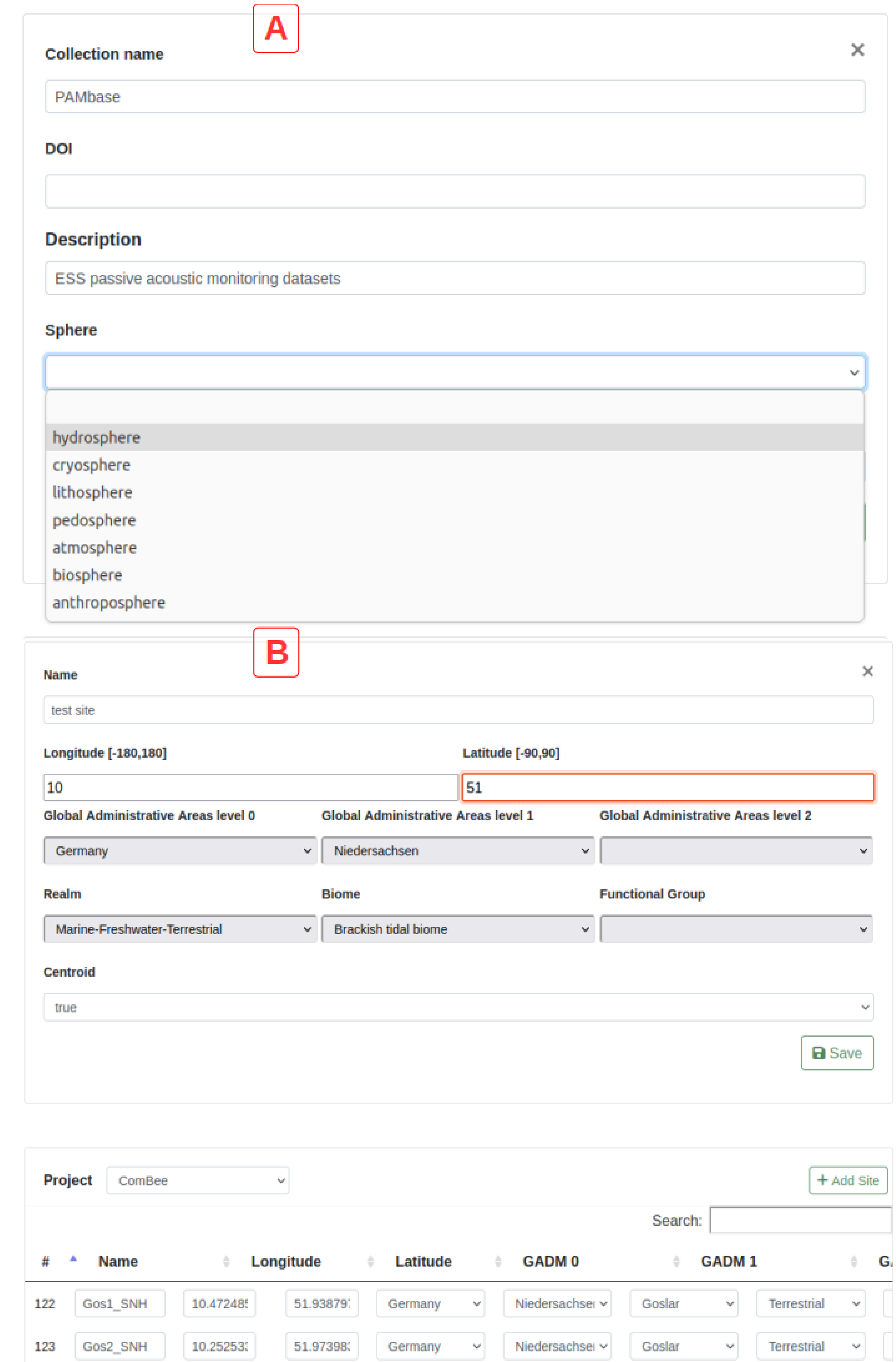

***Figure 7**: A) Collection creation mask. B) Site creation mask (top) and sites table (bottom).*

Soundscapes are also determined by the audio recording setup. While only a few guidelines exist for the terrestrial realm (Darras et al., 2018), actual standards exist for the acoustic monitoring meta-data in the marine realm (Stremmel and Struck, 2017). Audio recording parameters consist of the model and brand of the recorder, as well as the microphone used with that recorder. The microphone sensitivity (measured with a 1 kHz tone at 94 dB SPL) and signal-to-noise ratio (measured at 1 kHz as the interval between the self-noise and the same reference level of 94 dB SPL) are also fundamental specifications of the microphone with broad impacts on the recording of wildlife sounds (Darras et al., 2020a). Further audio settings such as frequency filters and gain settings are determined by the parameterisation of the recorder. We linked recordings to an expandable recorder table, which is itself linked to a microphone table, containing the most important fields related to audio quality and setup. Only compatible recorder-microphone combinations can be assigned to recordings, and this sensor information is now required when uploading recordings. Channel number is automatically extracted from the audio file using SoX[6] and stored. Similarly, recording durations are automatically derived from the uncompressed file's bit size, depth, and sampling frequency using SoX and logged. In the future, audio recorder settings (gain, frequency filters) should be included inside the recordings table.

Importantly, **raw data** should also be standardised. The calibration of sound recordings with data derived from sound level calibrators is an important feature that needs to be implemented to enable unbiased comparison of primary data across studies and application domains (Merchant et al., 2015). We strongly advocate that all acoustic sensors are calibrated to enable comparisons of sound levels across studies, on an absolute scale, to use sound pressure levels (dB SPL) instead of relative sound levels defined as units below zero, the maximum output of the corresponding system (dB). Absolute sound values enable comparisons between devices and studies. At the moment, ecoSound-web does not support standardised SPL from calibrated recordings but we plan to obtain absolute sound pressure level measures based on audio settings[7] and later, device calibrations.

Secondary data that are derived from defined analytical functions are straightforward to standardise. Mathematically-defined, standard acoustic indices can be computed for selected recordings. Similarly, mathematically Fast-Fourier-Transform derived spectrograms are generated with known window settings as determined by the user. Summarising recordings' data across the frequency axis into frequency spectra[8] is also a much-needed standard analysis feature. All these secondary data types need no further standardisation apart from the documentation of their calculations in the open-source code. However, to shift the focus from the analysis of natural sounds (biophony) to other types of sounds relevant in Earth system

[6] https://sox.sourceforge.net/
[7] https://github.com/ecomontec/ecoSound-web/issues/241
https://github.com/ecomontec/ecoSound-web/issues/257
https://github.com/ecomontec/ecoSound-web/issues/220
[8] https://github.com/ecomontec/ecoSound-web/issues/231

sciences, we implemented a sound typology based on a widely established classification to include the geophony and anthropophony (Pijanowski et al., 2011).

Sharing

To share data for scientific collaboration, they must be findable and accessible, according to the first two FAIR principles.

Meta-data can be shared through different channels. We recommend transparently publishing the database structure (Figure 2, (Darras et al., 2023)) and making the database itself openly available on established repositories. To that effect, the database underlying our main open project is regularly made available and updated on Zenodo (Darras, 2022), but other projects are closed. The meta-data of each table can also be downloaded as CSV files from the admin interface; downloads are subsets to the chosen project or collection. In the future, we would directly download SQL database packages where all the relevant collection or project data are included with tables linked explicitly with database relationships[9]. Conversely, we developed a feature to upload meta-data of recordings (so called “meta-recordings”) to quantify the acoustic monitoring effort of the vast majority of datasets that never get uploaded due to the unmanageably large volume of data that they represent.

We chose Creative Commons licenses to explicitly define and regulate the access of and usage rights for **raw data**. All Creative Commons licenses can be chosen during upload and edited anytime. The recordings themselves can be downloaded for offline use: any currently-generated audio file can be downloaded in compressed format (MP3 for audible sound, OGG for ultrasound). Findability can also be increased by linking raw data unambiguously to the corresponding publications and researcher profiles: We indexed collections and recordings with searchable DOIs, and recordists (i.e., users who uploaded the recording) can be identified with ORCID badges.

The **secondary data** generated from the raw data can easily be shared, but each type of data requires dedicated sharing functions. Furthermore, Fast-Fourier-Transform spectrograms can be downloaded (with a choice of selectable FFT window sizes) for any currently-generated view. The downloaded spectrogram file name includes the original audio file name as well as the coordinates in space and time. Further spectrogram parameters will be included in the future[10]. The annotations generated inside recordings can be shared with dedicated sharing links through HTTP GET requests that specify the recording identity as well as the coordinates in time and frequency. Larger downloads of the annotations contained within an entire collection can be made through the admin interface by any user, and the same applies to all the acoustic index computation results that were carried out by the corresponding user. Downloads only include the annotations that the currently logged-in user has access privileges to.

[9] https://github.com/ecomontec/ecoSound-web/issues/266
[10] https://github.com/ecomontec/ecoSound-web/issues/265

III. Challenges and Gaps

Most challenges were tackled by relying on and expanding the functionality of previously published open-source code. However, the design of the quality control, standardisation, and sharing guidelines and conceptualisation of how FAIR principles can be enforced in passive acoustic monitoring revealed how many efforts still need to be made, especially regarding the standardisation of the core, raw data (i.e., sound recordings) for truly re-usable data.

The second challenge lies in the fact that the part of the Earth System Science community that is using passive acoustic monitoring seems not to be networked yet, as shown in the current overview of platforms and tools in the introduction. We could therefore only reach out to the NFDI4Earth and NFDI4Biodiversity consortia for our user survey described in the following section, as we could not identify other networks to send the survey to. Thus, the “other stakeholders” and “partners [...] beyond NFDI4Earth” were not reached. This cultural challenge also made it difficult to find common ground within the Earth System Science community and to design recommendations that are more specific to this community.

IV. Relevance for the community and NFDI4Earth

Community

With the present pilot, the passive acoustic monitoring community clearly gained an improved software tool that implemented several FAIR principle-based recommendations, such as interoperability with ORCID, detailed sensors documentation, public recording collections, the integration of the spheres concept, a holistic sound typology, a meta-recording upload tool, and recording timeline views. On top of that, the community gained more fine-grained management of users, collections, and projects, as well as an interface for executing state-of-the-art deep learning models for the automatic ID of bird calls within soundscape recordings - the latter being especially valuable for ecoacousticians.

NFDI4Earth

Due to the nature of our pilot, we surveyed the needs of both NFDI4Earth and NFDI4Biodiversity researchers as users of passive acoustic monitoring tools. We proposed a classification for the acoustic data workflow (Darras et al., 2023) and designed an online questionnaire to compile general information about their studies, methods, measured variables, needs for software tools, and willingness to collaborate. The survey can be found here:
https://bildungsportal.sachsen.de/umfragen/limesurvey/index.php/975589?lang=en

We obtained only three full and fifteen partially-completed surveys. Only one respondent indicated to belong to NFDI4Earth and that person used active sampling methods for petrochemical investigations. A higher number of respondents (6) belonged to NFDI4Biodiversity, suggesting a stronger interest of that community. However, none of the respondents indicated to use passive sampling methods, which may indicate a misunderstanding (although we defined active and passive methods on the same page). Most of the following questions were left unanswered or incomplete. These humbling results may suggest either 1) a small interest of the NFDI4Earth community for developing common databases or tools for passive acoustic monitoring data; 2) a low awareness of the benefits of collaborative, interdisciplinary work and standardisation; 3) a too small sample size among the 60 NFDI4Earth partners for the sub-field of passive acoustic monitoring. In the future, the survey could be adapted and sent to larger networks.

V. Future Directions

Interoperability: APIs for servers and streaming devices

Ultimately, a successful integration of PAMbase and ecoSound-web into the Earth System Sciences will depend on their ability to communicate with a range of servers and recording devices through APIs (Application Programming Interfaces). Seamless integration of the vast amount of data that are stored in independent, more specialised platforms will bring more collaborators, expertise, and experience onboard. Continuously recording, and also streaming devices are more common in Earth System Sciences due to the less challenging, lower sampling frequencies (see Earthsound project) but are envisioned here for higher-frequency sampled data. A variety of platforms are also already established for biosphere monitoring, and the aim is not to supplant them, but to build data bridges between them. A first, future step would be to directly retrieve audio from URLs[11] and to link to reference call libraries[12].

Deep learning advances

We demonstrated the successful integration of Tensorflow neural networks to speed up soundscape annotation with artificial intelligence. Beyond the simple execution of algorithms, we need to develop new workflows and frontends that allow users to select existing manual annotations to fine-tune their own models. The back-end would select appropriate pre-trained models as a basis for the fine-tuning. The more user-friendly the process, the more likely the adoption of that feature will be, and the more rapidly we will be able to process the vast amounts of primary acoustic data that are not yet analysed with high accuracy, given the importance of contextual data (Lauha et al., 2022).

We explored the feasibility of human voice detection by scouting for the available, state-of-the art software tools. Voice activity detection enables the identification of speech regions in an audio signal, which is especially useful in the legal context, as privacy protection regulations often limit or forbid the use of acoustic monitoring methods when human voices are recorded without consent. As such, it could be used to mask parts of the recordings with high-pass filters to make them legally compliant. To cater to the increasing demand of voice activity detection,

[11] https://github.com/ecomontec/ecoSound-web/issues/218
[12] https://github.com/ecomontec/ecoSound-web/issues/333

also in other domains such as human speech processing, many open-source Python-based implementations are available on platforms such as GitHub. We compared five popular VAD implementations to assist users in choosing the most appropriate one for their specific applications (Table 1).

Table 1*: Overview of leading open-source algorithms for Voice activity detection*

Algorithm/ product name	**Taxonomic scope**	**Reference**	**Code**
VAD-python	Humans	NA	https://github.com/marsbroshok/VAD-python
Py-webrtcvad	Humans	NA	https://github.com/wiseman/py-webrtcvad
VadNet	Humans	http://www.isca-speech.org/archive/Interspeech_2018/abstracts/1238.html	https://github.com/hcmlab/vadnet
Voice Activity Detection project	Humans	NA	https://github.com/filippogiruzzi/voice_activity_detection
Pyannote.audio	Humans	https://arxiv.org/abs/1911.01255	https://github.com/pyannote/pyannote-audio

We recommend "Voice Activity Detection project" and Pyannote.audio as two recent deep learning-based solutions. Both implementations use TensorFlow and PyTorch, respectively, and offer pre-trained models, significantly reducing the development time and effort. PyTorch, a deep learning framework, is gaining popularity for developing state-of-the-art deep learning algorithms, however, it requires more expertise for developing the algorithms. Still, users may need to expand the post-processing labeling workflow for specific requirements (e.g., for merging consecutive detections, narrowing the frequency range, etc.). Notably, Pyannote.audio is the only implementation that is still being actively developed, which will likely enable longer support and enhanced features. Furthermore, we recommend to use VAD-python - which is not using state-of-the-art algorithms anymore - as a benchmark for evaluating more recent algorithms.

VI. References

Benoit-Bird, K.J., Lawson, G.L., 2016. Ecological Insights from Pelagic Habitats Acquired Using Active Acoustic Techniques. Annu. Rev. Mar. Sci. 8, 463–490. https://doi.org/10.1146/annurev-marine-122414-034001

Blickley, J.L., Patricelli, G.L., 2010. Impacts of Anthropogenic Noise on Wildlife: Research Priorities for the Development of Standards and Mitigation. J. Int. Wildl. Law Policy 13, 274–292. https://doi.org/10.1080/13880292.2010.524564

Boesch, R., Obrist, M., 2013. BatScope - Implementation of a BioAcoustic Taxon Identification Tool.

Cheng, F., Lindsey, N.J., Sobolevskaia, V., Dou, S., Freifeld, B., Wood, T., James, S.R., Wagner, A.M., Ajo-Franklin, J.B., 2022. Watching the Cryosphere Thaw: Seismic Monitoring of Permafrost Degradation Using Distributed Acoustic Sensing During a Controlled Heating Experiment. Geophys. Res. Lett. 49, e2021GL097195. https://doi.org/10.1029/2021GL097195

Darras, K.F.A., 2022. The Global Soundscapes Project: overview of datasets and meta-data. https://doi.org/10.5281/zenodo.6486836

Darras, K.F.A., Batáry, P., Furnas, B., Celis-Murillo, A., Wilgenburg, S.L.V., Mulyani, Y.A., Tscharntke, T., 2018. Comparing the sampling performance of sound recorders versus point counts in bird surveys: A meta-analysis. J. Appl. Ecol. 55, 2575–2586. https://doi.org/10.1111/1365-2664.13229

Darras, K.F.A., Batáry, P., Furnas, B.J., Grass, I., Mulyani, Y.A., Tscharntke, T., 2019. Autonomous sound recording outperforms human observation for sampling birds: a systematic map and user guide. Ecol. Appl. 29, e01954. https://doi.org/10.1002/eap.1954

Darras, K.F.A., Deppe, F., Fabian, Y., Kartono, A.P., Angulo, A., Kolbrek, B., Mulyani, Y.A., Prawiradilaga, D.M., 2020a. High microphone signal-to-noise ratio enhances acoustic sampling of wildlife. PeerJ 8, e9955. https://doi.org/10.7717/peerj.9955

Darras, K.F.A., Pérez, N., -, M., Hanf-Dressler, T., 2020b. BioSounds: an open-source, online platform for ecoacoustics. F1000Research 9, 1224. https://doi.org/10.12688/f1000research.26369.1

Darras, K.F.A., Pérez, N., Mauladi -, Dilong, L., Hanf-Dressler, T., Markolf, M., Wanger, T.C., 2023. ecoSound-web: an open-source, online platform for ecoacoustics. F1000Research. https://doi.org/10.12688/f1000research.26369.2

Darras, K.F.A., Pütz, P., Fahrurrozi, Rembold, K., Tscharntke, T., 2016. Measuring sound detection spaces for acoustic animal sampling and monitoring. Biol. Conserv. 201, 29–37. https://doi.org/10.1578/AM.39.1.2013.23

De Coensel, B., Sun, K., Botteldooren, D., 2017. Urban Soundscapes of the World: selection and reproduction of urban acoustic environments with soundscape in mind. INTER-NOISE NOISE-CON Congr. Conf. Proc. 255, 5407–5413.

Duarte, C.M., Chapuis, L., Collin, S.P., Costa, D.P., Devassy, R.P., Eguiluz, V.M., Erbe, C., Gordon, T.A.C., Halpern, B.S., Harding, H.R., Havlik, M.N., Meekan, M., Merchant, N.D., Miksis-Olds, J.L., Parsons, M., Predragovic, M., Radford, A.N., Radford, C.A., Simpson, S.D., Slabbekoorn, H., Staaterman, E., Opzeeland, I.C.V., Winderen, J., Zhang, X., Juanes, F., 2021. The soundscape of the Anthropocene ocean. Science 371. https://doi.org/10.1126/science.aba4658

Eisner, L., Duncan, P.M., Heigl, W.M., Keller, W.R., 2012. Uncertainties in passive seismic monitoring. Lead. Edge. https://doi.org/10.1190/1.3148403

Gallacher, S., Wilson, D., Fairbrass, A., Turmukhambetov, D., Firman, M., Kreitmayer, S., Mac Aodha, O., Brostow, G., Jones, K., 2021. Shazam for bats: Internet of Things for continuous real-time biodiversity monitoring. IET Smart Cities 3, 171–183.

https://doi.org/10.1049/smc2.12016
Gibb, R., Browning, E., Glover‐Kapfer, P., Jones, K.E., 2018. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. Methods Ecol. Evol. 10, 169–185. https://doi.org/10.1111/2041-210X.13101
Görföl, T., Huang, J.C.-C., Csorba, G., Győrössy, D., Estók, P., Kingston, T., Szabadi, K.L., McArthur, E., Senawi, J., Furey, N.M., Tu, V.T., Thong, V.D., Khan, F.A.A., Jinggong, E.R., Donnelly, M., Kumaran, J.V., Liu, J.-N., Chen, S.-F., Tuanmu, M.-N., Ho, Y.-Y., Chang, H.-C., Elias, N.-A., Abdullah, N.-I., Lim, L.-S., Squire, C.D., Zsebők, S., 2022. ChiroVox: a public library of bat calls. PeerJ 10, e12445. https://doi.org/10.7717/peerj.12445
Hedlin, M.A.H., Walker, K., Drob, D.P., de Groot-Hedlin, C.D., 2012. Infrasound: Connecting the Solid Earth, Oceans, and Atmosphere. Annu. Rev. Earth Planet. Sci. 40, 327–354. https://doi.org/10.1146/annurev-earth-042711-105508
Hupe, P., Ceranna, L., Le Pichon, A., Matoza, R.S., Mialle, P., 2022. International Monitoring System infrasound data products for atmospheric studies and civilian applications. Earth Syst. Sci. Data 14, 4201–4230. https://doi.org/10.5194/essd-14-4201-2022
Kahl, S., Wood, C.M., Eibl, M., Klinck, H., 2021. BirdNET: A deep learning solution for avian diversity monitoring. Ecol. Inform. 61, 101236. https://doi.org/10.1016/j.ecoinf.2021.101236
Keith, D.A., Ferrer-Paris, J.R., Nicholson, E., Bishop, M.J., Polidoro, B.A., Ramirez-Llodra, E., Tozer, M.G., Nel, J.L., Mac Nally, R., Gregr, E.J., Watermeyer, K.E., Essl, F., Faber-Langendoen, D., Franklin, J., Lehmann, C.E.R., Etter, A., Roux, D.J., Stark, J.S., Rowland, J.A., Brummitt, N.A., Fernandez-Arcaya, U.C., Suthers, I.M., Wiser, S.K., Donohue, I., Jackson, L.J., Pennington, R.T., Iliffe, T.M., Gerovasileiou, V., Giller, P., Robson, B.J., Pettorelli, N., Andrade, A., Lindgaard, A., Tahvanainen, T., Terauds, A., Chadwick, M.A., Murray, N.J., Moat, J., Pliscoff, P., Zager, I., Kingsford, R.T., 2022. A function-based typology for Earth's ecosystems. Nature 1–6. https://doi.org/10.1038/s41586-022-05318-4
Lauha, P., Somervuo, P., Lehikoinen, P., Geres, L., Richter, T., Seibold, S., Ovaskainen, O., 2022. Domain-specific neural networks improve automated bird sound recognition already with small amount of local data. Methods Ecol. Evol. 13, 2799–2810. https://doi.org/10.1111/2041-210X.14003
Liashchuk, O., Andrushchenko, Y., Zhukovskyi, V., Otruba, Y., Dovbysh, S., 2021. Infrasound and seismic monitoring of cryospheric events at the Vernadsky station region (West Antarctica). Presented at the 15th International Conference Monitoring of Geological Processes and Ecological Condition of the Environment, European Association of Geoscientists & Engineers, pp. 1–5. https://doi.org/10.3997/2214-4609.20215K2081
Merchant, N.D., Fristrup, K.M., Johnson, M.P., Tyack, P.L., Witt, M.J., Blondel, P., Parks, S.E., 2015. Measuring acoustic habitats. Methods Ecol. Evol. 6, 257–265. https://doi.org/10.1111/2041-210X.12330
Mohamed, A.-M.O., Paleologos, E.K., Howari, F.M., 2021. Chapter 19 - Noise pollution and its impact on human health and the environment, in: Mohamed, A.-M.O., Paleologos, E.K., Howari, F.M. (Eds.), Pollution Assessment for Sustainable Practices in Applied Sciences and Engineering. Butterworth-Heinemann, pp. 975–1026. https://doi.org/10.1016/B978-0-12-809582-9.00019-0
Nikolaev, A.V., Belyakov, A.S., Lavrov, V.S., Zhigalin, A.D., 2005. Geoacoustic monitoring as a means for investigating the state of the lithosphere and for earthquake forecasting. Acoust. Phys. 51, S122–S130. https://doi.org/10.1134/1.2133960
Parsons, S., Jones, G., 2000. Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks. J. Exp. Biol. 203, 2641–2656. https://doi.org/10.1242/jeb.203.17.2641

Pijanowski, B.C., Farina, A., Gage, S.H., Dumyahn, S.L., Krause, B.L., 2011. What is soundscape ecology? An introduction and overview of an emerging new science. Landsc. Ecol. 26, 1213–1232. https://doi.org/10.1007/s10980-011-9600-8

Priyadarshani, N., Marsland, S., Castro, I., 2018. Automated birdsong recognition in complex acoustic environments: a review. J. Avian Biol. 49, jav-01447. https://doi.org/10.1111/jav.01447

Redgwell, R.D., Szewczak, J.M., Jones, G., Parsons, S., 2009. Classification of Echolocation Calls from 14 Species of Bat by Support Vector Machines and Ensembles of Neural Networks. Algorithms 2, 907–924. https://doi.org/10.3390/a2030907

Rost, S., Thomas, C., 2002. Array Seismology: Methods and Applications. Rev. Geophys. 40, 2-1-2–27. https://doi.org/10.1029/2000RG000100

SoX - Sound eXchange | HomePage [WWW Document], n.d. URL http://sox.sourceforge.net/ (accessed 8.24.20).

Steffen, W., Richardson, K., Rockström, J., Schellnhuber, H.J., Dube, O.P., Dutreuil, S., Lenton, T.M., Lubchenco, J., 2020. The emergence and evolution of Earth System Science. Nat. Rev. Earth Environ. 1, 54–63. https://doi.org/10.1038/s43017-019-0005-6

Stremmel, N.B., Struck, C.J., 2017. Acoustical standards news. J. Acoust. Soc. Am. 142, 631–640. https://doi.org/10.1121/1.4996107

The Earthsound Project [WWW Document], n.d. URL https://www.earthsound.earth/index.html (accessed 11.2.22).

Thomisch, K., Flau, M., Heß, R., Traumueller, A., Boebel, O., 2021. OPUS - An Open Portal to Underwater Soundscapes to explore and study sound in the global ocean, in: EPIC35th Data Science Symposium, Virtual Meeting, 2021-01-22-2021-01-22. Presented at the 5th Data Science Symposium, virtual meeting.

Ulloa, J.S., Haupert, S., Latorre, J.F., Aubin, T., Sueur, J., 2021. scikit-maad: An open-source and modular toolbox for quantitative soundscape analysis in Python. Methods Ecol. Evol. 12, 2334–2340. https://doi.org/10.1111/2041-210X.13711

Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Santos, L.B. da S., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., Hoen, P.A.C. 't, Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., Schaik, R. van, Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., Lei, J. van der, Mulligen, E. van, Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data. https://doi.org/10.1038/sdata.2016.18

Xeno-canto Foundation, 2012. Xeno-canto: Sharing bird sounds from around the world. Xeno-canto Foundation Amsterdam.