



SoundLib: An Underwater Sound Library of the North Sea


Bram Cuyx, Clea Parcerisas, Julia Aubach, Martin Jälmy, Jorge Mendes, Franziska Regel, and Elisabeth Debusschere

Contents

Introduction	2
Background and Motivation	2
Objectives	3
Existing Datasets and Repositories	4
Methods	5
Data Collection	5
Data Contribution	6
Data Presentation	8
Software Development Kit (SDK)	10
Conclusion	11
References	11

Abstract

Passive acoustic monitoring (PAM) and soundscape analysis provide powerful tools to unlock biodiversity information embedded in underwater sound. These noninvasive techniques are gaining traction among marine bioacousticians for

OPEN ACCESS with major support from 

B. Cuyx (✉) · C. Parcerisas · J. Aubach · E. Debusschere
Flanders Marine Institute (VLIZ), Marine Observation Centre, Oostende, Belgium
e-mail: bram.cuyx@vliz.be; clea.parcerisas@vliz.be; julia.aubach@vliz.be;
elisabeth.debusschere@vliz.be

M. Jälmy
Department of Informatics, Universität Hamburg, Hamburg, Germany
e-mail: martin.jalmy@uni-hamburg.de

J. Mendes · F. Regel
Flanders Marine Institute (VLIZ), Data Centre, Oostende, Belgium
e-mail: jorge.mendes@vliz.be; franziska.regel@vliz.be

their effectiveness in species monitoring. The authors present the SoundLib project, an initiative to build a comprehensive library of acoustic events and long-term underwater sound trends from the North Sea.

Sound plays an important role in aquatic communication and as such can be used to assess the health of marine ecosystems by analyzing the soundscape in terms of biophony, geophony, and anthrophony. To contribute to this understanding, the authors employ PAM in the Belgian part of the North Sea (BPNS) and contribute a dataset of annotated long-term recordings.

To facilitate broader access and collaboration, the authors have developed an application programming interface (API) and a Python-based software development kit (SDK) for SoundLib. The SoundLib application allows researchers to query the recordings and its metadata, which includes annotations for detected events. Users can also contribute audio recordings, both long-term and isolated sound events, which can be processed using the provided standardized methods. By leveraging this SDK, researchers can efficiently analyze marine soundscapes and contribute to a growing repository of acoustic data.

Keywords

Passive acoustic monitoring · Machine learning · Sound event detection · Bioacoustics · Soundscapes · Dataset

Introduction

Background and Motivation

As North Sea exploitation becomes increasingly prominent, sound levels rise quickly while soundscapes change. Understanding how human activity affects local marine ecosystems is essential to mitigate adverse effects arising from this increased exploitation. (Jalkanen et al. 2022), (Popper and Hawkins 2019). Underwater sound is a crucial component of ecosystem health; as such, it must be monitored and understood (Erbe et al. 2022) (European Commission 2017). As the Belgian part of the North Sea (BPNS) is a highly turbid and dynamic environment, the line-of-sight can be limited to a few centimeters, making visual analysis of the ecosystem restrictive and incomplete. Contrary to this, sound carries farther and faster underwater than through air due to the medium's higher density (Erbe et al. 2022). A single deployed hydrophone could detect the presence of a shipping lane kilometers away. Sound is also an important modality for many marine species for communication or to obtain information about their environment (Popper and Hawkins 2019). As exploitation of the North Sea intensifies, this modality is increasingly being disrupted, restricting the information that these species can convey and receive (Duarte et al. 2021).

Underwater noise levels and soundscape can be monitored using passive acoustic monitoring (PAM). In PAM, hydrophones can continuously record sounds over

extended periods; these can then either be collected periodically or analyzed in real time. These recordings are used to describe the soundscape, as defined in *ISO 18405: 2017 Underwater Acoustics Terminology*, 2017, as: “the characterization of ambient sound in terms of its spatial, temporal, and frequency attributes, and the types of sources contributing to the sound field.” Analyzing this soundscape can help study how human activity affects marine ecosystems.

A first method for analyzing the soundscape is its description in terms of long-term spectral information, providing an evolution of sound pressure levels (SPL) in separate frequency bands. It is important to note that particle velocity might be a more relevant metric to monitor, but the process of doing so is more complex and expensive; thus, these standardized approaches focus on SPL instead. Although this does not provide detailed information on the origin of recorded sounds, it is useful to monitor noise pollution over the long term (Merchant et al. 2022).

A second approach is to try to identify the origin of the sounds comprising the soundscape by analyzing it into its constituent sound events by means of sound event detection (SED) (Mesaros et al. 2021). Individual sound events are detected before being categorized according to their origin, either biophonic, geophonic, or anthroponic. Biophonic sound events could be the click of a dolphin or the vocalization of a fish. The sounds of rain, waves, and movement of sand banks are categorized as geophony. Any sound originating from human activities, e.g., a diver, a boat, or the sound of a wind farm, is categorized as anthrophony. This is the most general classification; within each of these categories a more granular class can be assigned to each sound, e.g., which species produced the sound and even for what purpose, or what type of boat passes by. As the classifications get more precise, they get more challenging to assign. To best represent all sounds, even those of unknown origin, working with these multiple levels of classification is useful and necessary.

Objectives

The Marine Sound Library (SoundLib) functions as a collaborative library of underwater sound data in the North Sea. Serving three distinct audiences identified as the general public, researchers, and policymakers. SoundLib is presented in two distinct ways: the web platform and the Python software development kit (SDK). Each of these serves its intended audience.

Awareness should be created in the general public as to the importance of the effect sound has on marine ecosystems and the problem of noise pollution. Through visualization of sound level trends over the years per station and across stations and through sound event visualization and auditory display, the authors introduce the general public to this hidden world and educate them on the importance of sound levels, sound sources, and vocalizations of marine fauna.

Marine researchers can contribute and make use of the data on the SoundLib platform. One way to achieve this is by introducing a metadata standard, ensuring data is well documented, traceable, and linked to its original contributor as well as providing standard scripts to process the data according to internationally agreed

parameters (e.g., hybrid millidecade spectra as defined in ISO 7605:2025). This allows users to verify the data's authenticity and reliability and compare it to other measurements. By adding data collected on a platform that is user-friendly, data also becomes more citable and reusable.

SoundLib aims to grow into a comprehensive and valuable resource for the marine bioacoustics community by allowing easy contribution of sound events, long-term spectral information, associated metadata, and new annotations for existing events. This will support the effort of identifying the different soniferous species in the North Sea soundscape, other natural sources, and the sources that pressure their ecosystems. By providing many repetitions of each sound class, SoundLib can also serve as a resource for training SED algorithms.

A sound might be found in multiple recordings and recognized without knowing its exact origin. This can be an important opportunity for collaboration; comparing an unknown sound to a dataset of identified ones can help narrow down its possible sources. Collaboration should be facilitated to have more complete descriptions of each marine region's ecosystems.

To policymakers, SoundLib should function as a source to inform evidence-based marine exploitation policies and marine spatial planning. This can be achieved by incorporating metadata on the sensors used for measurements, containing details on calibration date and resulting (frequency-dependent) sensitivity, ensuring the SPLs are correctly determined from recordings. The long-term trends in underwater sound levels can be visualized in the Marine Strategy Framework Directive (MSFD) GES D11 (European Commission 2017).

Existing Datasets and Repositories

Several repositories for passive acoustics data already exist, each of them with its own focus.

FishSounds (Looby et al. 2023) and DOSITS (University of Rhode Island 2025) are educational resources collecting sounds and descriptions of the different soniferous species of fish or providing information on underwater sound.

GLUBS (Parsons et al. 2022) is a platform that aims to facilitate the sharing of underwater sound data among researchers. It provides an overview of underwater acoustic datasets in the world and existing packages to process these. However, there is no guarantee that the data is formatted in a similar way or that the metadata is consistent. This makes it more challenging to use this as a single dataset for training SED algorithms.

ICES (International Council for the Exploration of the Sea 2025) provides a repository of continuous underwater noise data focused on reporting SPL to support HELCOM/OSPAR underwater noise monitoring strategies.

Each repository on this nonexhaustive list is a valuable resource for marine bioacousticians. SoundLib aims to combine some of the strengths of each of these by providing an extensive and well-curated dataset of underwater sounds while providing a metadata standard that allows external contributions to be easily

integrated and utilized. By including both SPL reporting and annotated sound events, SoundLib provides a unique resource for studying the effects of noise levels on the presence of different sound classes, helping assess the effectiveness of OSPAR/HELCOM measures to reduce underwater radiated noise.

Methods

Data Collection

The two platforms used for the provided data collection are shown in Fig. 1, a multi-use platform, described by Calonge et al. (2024), in Fig. 1 (1), and a seabed lander in Fig. 1 (2). These are examples of platforms that can be used for underwater data collection.

These platforms are deployed on the seabed for a period typically lasting 3 months. For PAM, these moorings are equipped with at least a hydrophone, recorder, and an acoustic release. Additional sensors collecting other information, or multiple hydrophones, can be added to the mooring depending on the research objectives. After a deployment, the data from the hydrophones is downloaded. The data is processed in two distinct ways. Firstly, long-term spectral information is extracted in the format of hybrid millidecade bands (HMB) as defined in ISO 7605:2025. The Monterey Bay Aquarium Research Institute (MBARI) in collaboration with VLIZ has developed a Python package called PyPAM (Parcerisas 2023)-Based Processing (PBP) (MBARI 2025) to efficiently calculate HMBs from long-term recordings.

Secondly, the soundscape is analyzed in terms of its comprising events. This is achieved using machine learning (ML) methods, where a recording is cut into overlapping frames of a predefined length. These frames are then fed into a network

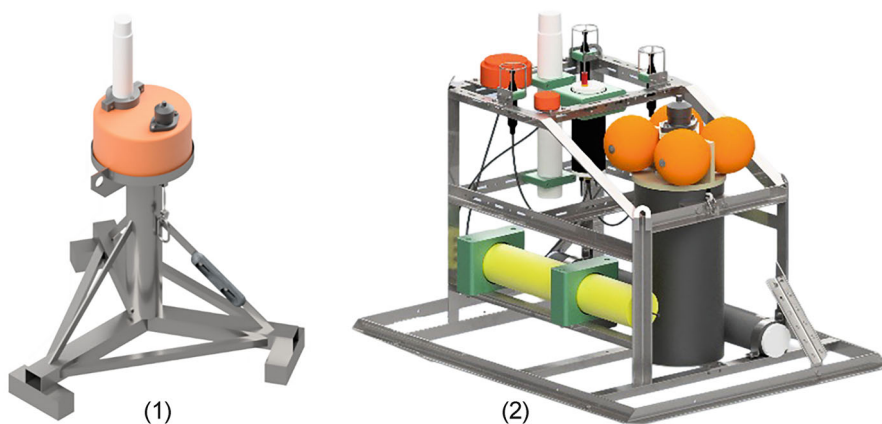


Fig. 1 VLIZ multi-use platform (Calonge et al. 2024)

which generates an embedding, a compressed representation of the information in that specific frame. For the contributed dataset, the pretrained SurfPerch model is used to create the embeddings. This embedding is then fed into a classifying network, a much simpler network than the embedding network; within a few specified classes, this network can then predict if a specific class is present in a frame (Williams et al. 2025).

The human-annotated labels that form the basis of the SoundLib dataset and used as input for such networks have been created with Raven Pro (K. Lisa Yang Center for Conservation Bioacoustics at the Cornell Lab of Ornithology 2024). The labeling process follows an annotation protocol that provides a set of systematic guidelines for both general and sound-type-specific labeling to ensure consistency within and across sound classes. An integral part of the annotation process is the adherence to strict naming conventions. The protocol specifies that a sound signature is assigned a general classification (biophony, geophony, anthrophony, or unknown) and a source, which represents the most detailed level of its known origin (e.g., pile driving, fish vocalization, dolphin whistle, unknown). Additionally, each sound signature is assigned a descriptive label, either as a phonetic representation (onomatopoeia) that best captures it (e.g., clack, pop, gulp) or a name of a source that makes a similar sound (e.g., jingle bell, jackhammer, drumroll). With this, acoustically salient and relevant sounds are registered in a hierarchical and detailed manner. By having many repetitions for each class, recorded at different times and locations, the dataset can help improve the generalization of the trained models.

Data Contribution

Metadata Standard

SoundLib is open for contribution of external datasets. A metadata standard, linked on <https://marinesoundlib.org/data/en/data-contribution>, is used to keep track of the information regarding each deployment configuration, including specifications of hydrophones and stations. Furthermore, a deployment is the period between the installation of a measurement setup and its retrieval (it may also be a transect along a specified track) and can contain multiple continuous recordings. Each recording is a continuous sequence of audio files without gaps (i.e., no unsampled intervals) and is associated with a specific recorder and hydrophone configuration; for every recording the associated files, sound events, embeddings, and annotations are enumerated. Sound events are understood as specific segments of audio that contain a sound of interest; annotations link a sound event to a specific class or label, provided by a human annotator or machine learning model. For every sound event, multiple embeddings can be stored, generated by different models.

Additional annotations can be made for existing events; these do not overwrite existing annotations but are added to the list of annotations for that event, allowing multiple experts to add their own interpretations. This can be seen in Table 5 in the metadata standard under `events[].annotations[]`, showing for every event a list of

Library / Data contribution / Contribute / Deployments / Create

BC

Name * ⓘ

Dataset * ⓘ

Author * ⓘ

Licence type * ⓘ

Time - Ship

Location details

Comments ⓘ

Cancel Save

* indicates a mandatory field

Fig. 2 Metadata contribution tool on the Marine SoundLib website

annotations. Every annotation contains information about the contributor, date, method of classification, and the confidence with which the label is assigned.

A web-based metadata contribution tool is provided, as seen in Fig. 2, simplifying the process of metadata generation. This figure shows how metadata for a deployment can be entered in a user-friendly way; similar ones exist for recorder, hydrophone, and detections.

FAIR Data and Licenses

Marine scientists are encouraged to upload long-term recordings to the platform. Authors and institutes can be identified through the MarineInfo platform (<http://marineinfo.org>), where any marine scientist can create an account and be identified through a unique identifier. Whenever the contributed data is queried, these attributions to author, institute, and dataset are returned alongside the data. For full deployments, within the recordings, only links to the audio files need to be provided, such that each institute retains and manages its own data and only the SPL trends can be downloaded by external scientists. For sound events, on the other hand, the raw audio is saved to SoundLib, such that people who want to train models can access it.

To ensure that the data contributed to SoundLib adheres to the FAIR principles (Wilkinson et al. 2016), several measures are taken. Data is stored in a structured database with unique identifiers for each entity (deployment, recorder, hydrophone, event, annotation) and can be queried according to multiple parameters, e.g., number of channels, locations, sound types, and others. This ensures that data is findable and accessible. Interoperability is ensured by using standard formats and provided protocols to process data. Contributors are encouraged to choose appropriate licenses for contributed data.

Data Presentation

Users are welcomed to the website with a map of all stations, and data is collected at and uploaded to the database as shown in the screenshot in Fig. 3. By clicking on a station, users can access the data collected there. This is presented as an HMB plot, as seen in Fig. 4, showing the evolution of the SPL over time for a specific station. Users can zoom in on specific time periods to get a more detailed view of the data and explore the events that were detected at that location. Figure 5 shows how the SPL for different stations can be compared in specific frequency ranges.

The page with the sound events shown in Fig. 6 allows users to explore the annotated sound events. This list can be filtered by different criteria such as class, name, and station. By clicking on a specific event, users can access more information, such as metadata, audio clip, spectrogram, or download the sound.

A visualization of these event embeddings is presented in Fig. 7, where every event is presented as a dot, and clusters of sounds appear where the embeddings lie closest together. This visualization is made using UMAP (McInnes et al. 2020), a dimensionality reduction technique, projecting the embeddings onto a 2D space.

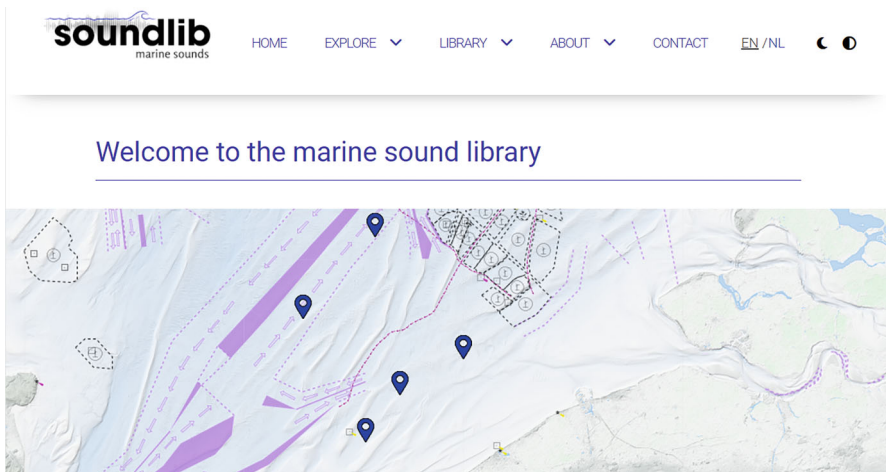


Fig. 3 Welcome page of Marine SoundLib with interactive map

Explore / Stations / Station details

bpns-Grafton

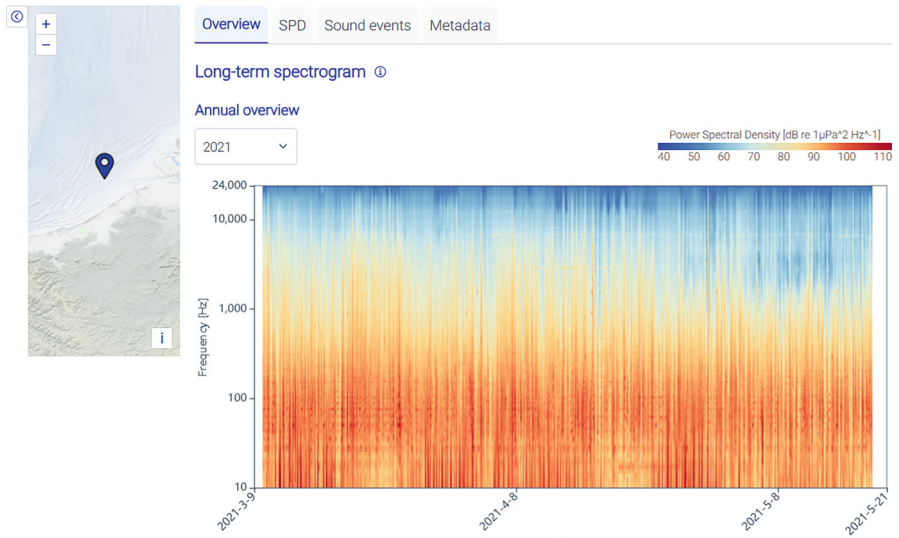


Fig. 4 Investigate the evolution of sound pressure levels as a function of frequency for a station over a deployment



Fig. 5 Compare the sound pressure levels for different stations, represented by different colors, over time for specified frequency ranges

When a sound is added to the database, its embedding is added to the plot allowing users to see how similar the new sound is to existing ones in the database; if it shows up within an existing cluster, it might indicate the sound belongs to that same class. When more annotated data is added to this dataset, the origin of unknown sounds can be more easily identified by comparing them to the existing sounds in the database.

get access to all the data regarding the sensor that was used and its calibration values and date, allowing users to process the audio from floating point values to pressure values. Within the SDK, scripts will be included to standardize the processing of deployments, e.g., calculating HMBs and uploading these to the SoundLib database, or for calculating and uploading embeddings for detected events along with all necessary metadata. This will facilitate the contribution of data to the platform, ensuring data is processed in a consistent manner.

Conclusion

The authors present a platform designed to facilitate collaboration and curation of underwater passive acoustic data, both in the form of long-term sound trends and annotated events. By centralizing acoustic data in a single repository, data usability is enhanced, making for a more efficient use of resources and expertise in the field.

To support accurate representation and interoperability of different measurement procedures, the authors present a metadata standard. This standard includes information such as sensor model and sensitivity, enabling contributed data to be used reliably in assessing underwater noise pollution across different locations. It also ensures proper attribution, incentivizing collaboration.

Annotations are presented for four separate classes of sound: anthrophony, geophony, biophony, and unknown. Within these classes, more granular classes can be assigned to each event. The long-term recordings are presented with their hybrid millidecade bands spectral information, providing a precise representation of the general state of noise pollution in the North Sea.

SoundLib is presented in two distinct ways: the web platform and the Python software development kit. While the web platform is useful for creating awareness in the general public and can function as an information source for policymakers, researchers can interact with the data using a software development kit, allowing them to directly query data from the database and contribute their own. In conclusion, the SoundLib stimulates collaboration and accelerates the field of underwater soundscapes and eco-acoustics.

Competing Interest Declaration The authors have no competing interests to declare that are relevant to the content of this manuscript.

References

- Calonge A, Navarro González P, Mortelmans J et al (2024) Cetacean passive acoustic network in the Belgian part of the North sea. *Sci Data* 11:971. <https://doi.org/10.1038/s41597-024-03806-y>
- Duarte CM, Chapuis L, Collin SP et al (2021) The soundscape of the anthropocene ocean. *Science* 371:eaba4658. <https://doi.org/10.1126/science.aba4658>
- Erbe C, Duncan A, Vigness-Raposa KJ (2022) Introduction to sound propagation under water. In: Erbe C, Thomas JA (eds) *Exploring animal behavior through sound: volume 1: methods*. Springer International Publishing, Cham, pp 185–216

- European Commission (2017) Commission Decision (EU) 2017/848
- International Council for the Exploration of the Sea (ICES) (2025) Continuous underwater noise
- ISO (2017) ISO 18405:2017 Underwater acoustics terminology
- ISO (2025) ISO 7605:2025 Underwater acoustics – measurement of underwater ambient sound
- Jalkanen J-P, Johansson L, Andersson MH et al (2022) Underwater noise emissions from ships during 2014–2020. *Environ Pollut* 311:119766. <https://doi.org/10.1016/j.envpol.2022.119766>
- K. Lisa Yang Center for Conservation Bioacoustics at the Cornell Lab of Ornithology (2024) Raven Pro: interactive sound analysis software
- Looby A, Vela S, Cox K et al (2023) FishSounds version 1.0: a website for the compilation of fish sound production information and recordings. *Eco Inform* 74:101953. <https://doi.org/10.1016/j.ecoinf.2022.101953>
- MBARI (2025) MBARI-PBP: PyPAM-based data processing for ocean audio archives
- McInnes L, Healy J, Melville J (2020) UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction
- Merchant ND, Putland RL, André M et al (2022) A decade of underwater noise research in support of the European Marine Strategy Framework Directive. *Ocean Coast Manag* 228:106299. <https://doi.org/10.1016/j.ocecoaman.2022.106299>
- Mesaros A, Heittola T, Virtanen T, Plumbley MD (2021) Sound Event Detection: A tutorial. *IEEE Signal Process Mag* 38:67–83. <https://doi.org/10.1109/MSP.2021.3090678>
- Parcerisas C (2023) PyPAM: a package to process underwater acoustics time series
- Parsons MJG, Lin T-H, Mooney TA, et al (2022) Sounding the call for a global library of underwater biological sounds. *Front Ecol Evol* 10. <https://doi.org/10.3389/fevo.2022.810156>
- Popper AN, Hawkins AD (2019) An overview of fish bioacoustics and the impacts of anthropogenic sounds on fishes. *J Fish Biol* 94:692–713. <https://doi.org/10.1111/jfb.13948>
- University of Rhode Island (2025) Discovery of Sound in the Sea (DOSITS)
- Wilkinson MD, Dumontier M, Aalbersberg IJ et al (2016) The FAIR guiding principles for scientific data management and stewardship. *Sci Data* 3:160018. <https://doi.org/10.1038/sdata.2016.18>
- Williams B, van Merriënboer B, Dumoulin V et al (2025) Using tropical reef, bird and unrelated sounds for superior transfer learning in marine bioacoustics. *Philos Trans R Soc B Biol Sci* 380: 20240280. <https://doi.org/10.1098/rstb.2024.0280>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

