

# Bottom-Up Research Data Repositories

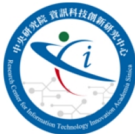
SciDataCon @ International Data Week 2022  
June 21, 2022

Tyng-Ruey Chuang

with Cheng-Jen Lee, Chia-Hsun Wang, and Ming-Syuan Ho

[data.contact@depositor.io](mailto:data.contact@depositor.io)

Institute of Information Science &  
Research Center for Information Technology Innovation  
Academia Sinica, Taiwan

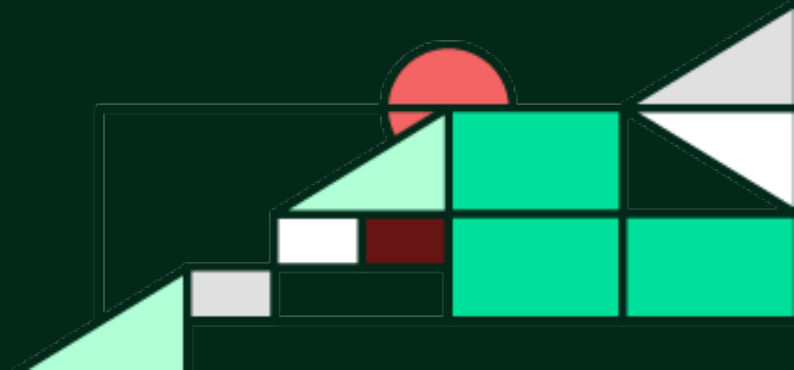


<https://data.depositor.io/en/dataset/scidatacon2022>



# An Open Repository for All

- The *depositar* – How does it work?
- A bottom-up data repository: Process and tools
- Discussion
  
- A continuation of a presentation at Open Repositories 2021
  - *Experience in Moving Toward An Open Repository For All*  
by Tyng-Ruey Chuang, Cheng-Jen Lee, Chia-Hsun Wang, and Yu-Huang Wang



The screenshot shows the Data Depositor interface for the dataset 'Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan'. The page is structured as follows:

- Header:** Depositor logo, navigation links (Datasets, Topics, Projects, About, Help), and user options (Log In, Register, @2).
- Project Sidebar:**
  - Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan:** Project title and description.
  - Followers:** 0.
  - Project:** Ocean Biodiversity Listening Project.
  - Social:** Twitter and Facebook links.
  - License:** CC-BY 4.0.
  - Cite as:** American Psychol.
  - Dataset extent:** Map showing the location of Sesoko Island, Okinawa, Japan.
  - Other Access:** Information on how to access the dataset via various methods.
- Main Content Area:**
  - Dataset:** Overview of the dataset, including a description and recording locations.
  - Recording Locations:** Details of the three long-term recording sites (Site A, Site B, Site C).
  - Acoustic Recorders:** Information about the AUSOMS-mini stereo recorders used.
  - Configuration of Audio Recording:** Technical details of the recording setup.
  - Field Deployment:** Information about the deployment of the recorders.
  - Data Processing:** Details of the audio processing pipeline.
  - Associated Publication:** Reference to the scientific paper by Tzu-Hao Lin et al. (2021).
  - Data and Resources:** Links to audio data, long-term spectrograms for each site, and codes for data access and analysis.
  - Tags:** Acoustic diversity, Acoustic habitat, Coral reef, Mesophotic corals, Noise, Ocean sound, Remote sensing, Underwater soundscape.
  - Wikidata Keywords:** soundscape, coral reef.
  - Basic Information:**
    - Data Type:** Source code, Audiovisual data, Scientific and statistical data formats.
    - Language:** English (eng).
    - Spatio-temporal Information:**
      - Temporal Resolution: Daily
      - Start Time: 2017-05
      - End Time: 2018-07
      - Spatial Coverage: X.min: 127.8553390572779, X.max: 127.88097380893306, Y.min: 26.630362980584657, Y.max: 26.68047930832328
  - Management Information:**
    - Author:** Tzu-Hao Lin, Tomonari Akamatsu, Frederic Sinniger, Saki Hari
    - Contact Person:** Tzu-Hao Lin (schorikopl@gmail.com)

# A Sample Dataset

<https://data.depositor.io/en/dataset/coral-reef-sesoko>

## A tour of the dataset

- Long description of the dataset and the project
- Data and (external) resources
- Tags and Wikidata keywords
- Basic information
- Spatio-temporal information
- Management information
- License
- Citation snippet
- Data endpoints
  - JSON-API
  - RDF serializations

## Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan

Followers



0

### Project



### Ocean Biodiversity Listening Project

Project Website The ocean is full of sounds that are generated from geophysical events, marine animals, and human activities. By using a hydrophone (a microphone for underwater... [read more](#)

Dataset  Topics  Activity Stream  History

## Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan



This dataset is an archive of audio data of shallow-water and upper-mesophotic coral reefs off Sesoko Island, Okinawa, Japan. Python codes to visualize the audio data were also provided in a notebook based on Google Colab.

### Recording Locations

Three long-term recording sites were established since May 2017. Site A (N26.635° E127.865° ) is located on the southeast coast of Sesoko Island and in front of the Sesoko Station of the University of the Ryukyus. The water depth is 1.5 m. Site B (N26.665° E127.869° ) is located at the bottom of a reef slope on the north of Sesoko Island and the west of Toguchi Port. The water depth is 20 m. Site C (N26.670° E127.866° ) is located on a nearly flat plateau to the north of Sesoko Island and the west of Toguchi Port. The water depth is 40 m.

### Acoustic Recorders

[AUSOMS-mini stereo recorders](#) (AquaSound, Kobe, Japan) were used to collect underwater sounds. From May 2017 to July 2018, six AUSOMS-mini recorders were used: 14-0106, 14-0107, 15-0106, 15-0107, 15-0109, 15-0110.

### Configuration of Audio Recording

(1) Duty Cycle: continuous. (2) Sampling Rate: 44.1kHz. (3) Channels: 2. (4) File Format: MP3 (128 kbps). (5) Audio Gain: High. (6) High Pass Filter: Off.

### Field Deployment

At each recording site, one AUSOMS-mini stereo recorder was fixed to a cement





**Social**

---

Twitter

---

Facebook

---

**License**

---

[CC-BY 4.0](#) [OPEN DATA](#)

---

**Cite as** Beta

---

Tzu-Hao Lin, Tomonari Akamatsu, Frederic Sinniger, Saki Harii. (2021). *Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan* (Version 2021-01-09T09:11:31.023608) [Data set]. Retrieved from <https://data.depositar.io/en/dataset/coral-reef-sesoko>

**Cut to clipboard**

Deployment and recovery of recorders were conducted by divers.

**Data Processing**

Audio recordings generated by AUSOMS-mini recorders were saved in MP3 format. Each MP3 is about 8-hour long and do not have a time stamp on the file name. To facilitate data management, we segmented the 8-hour long MP3 into WAV files of 5-min duration.

We used the [LTSA\\_gui](#) to generate long-term spectrograms (LTS) and save the LTS in mat files. Each mat file contains median-based LTS and mean-based LTS. Median-based LTS was obtained by measuring median power spectral densities within each 5-min segment. Mean-based LTS was obtained by measuring mean power spectral densities within each 5-min segment.

**Associated Publication**

Tzu-Hao Lin, Tomonari Akamatsu, Frederic Sinniger, Saki Harii (2020) Exploring coral reef biodiversity via underwater soundscapes. [Biological Conservation, 253: 108901.](#)

**Data and Resources**

	<a href="#">Audio data</a>	<a href="#">Explore</a>
A link to a shared Drive folder of underwater recordings (WAV) and long-term...		
	<a href="#">Long-term spectrogram of Site A</a>	<a href="#">Explore</a>
A mat file contains the median- and mean-based long-term spectrograms.		
	<a href="#">Long-term spectrogram of Site B</a>	<a href="#">Explore</a>
A mat file contains the median- and mean-based long-term spectrograms.		
	<a href="#">Long-term spectrogram of Site C</a>	<a href="#">Explore</a>
A mat file contains the median- and mean-based long-term spectrograms.		





Map tiles & Data by OpenStreetMap [↗](#),  
under CC BY-SA [↗](#).

### Other Access

The information on this page (the dataset metadata) is also available in these formats:

[</>JSON-API](#)

### [RDF serializations](#)

based on DCAT 2: **Beta**

[</>JSON-LD](#)

[</>Turtle](#)

[</>XML](#)

via the [CKAN API](#) [↗](#)



### Tags

- Acoustic diversity
- Acoustic habitat
- Coral reef
- Mesophotic corals
- Noise
- Ocean sound
- Remote sensing
- Underwater soundscape

### Wikidata Keywords

- soundscape
- coral reef

### Basic Information

<b>Data Type</b>	<ul style="list-style-type: none"> <li>Source code</li> <li>Audiovisual data</li> <li>Scientific and statistical data formats</li> </ul>
<b>Language</b>	English (eng)

### Spatio-temporal Information

<b>Temporal Resolution</b>	Daily
<b>Start Time</b>	2017-05
<b>End Time</b>	2018-07
<b>Spatial Coverage</b>	<a href="#">show more</a>
<b>X.min</b>	127.8553390572779
<b>X.max</b>	127.88097380893306
<b>Y.min</b>	26.630362980584657
<b>Y.max</b>	26.68047930832328

### Management Information

<b>Author</b>	Tzu-Hao Lin, Tomonari Akamatsu, Frederic Sinniger, Saki Harii
<b>Contact Person</b>	Tzu-Hao Lin





Map tiles & Data by OpenStreetMap  
under CC BY-SA



### 其他存取方式

此頁面上的資訊 (資料集之後設資料) 也提供以下格式：

</>JSON-API

RDF 序列化輸出 (修改自 DCAT 2) : Beta

</>JSON-LD </>Turtle </>XML

經由 [CKAN API](#)



### 標籤

- Acoustic diversity
- Acoustic habitat
- Coral reef
- Mesophotic corals
- Noise
- Ocean sound
- Remote sensing
- Underwater soundscape

### Wikidata 關鍵字

- 聲景
- 珊瑚礁

### 基本資訊

資料類型	<ul style="list-style-type: none"> <li>原始碼</li> <li>影音資料</li> <li>科學與統計資料</li> </ul>
語言	英文 (eng)

### 時空資訊


時間解析度	日
起始時間	2017-05
結束時間	2018-07
空間範圍	<a href="#">顯示更多</a>
空間範圍.X.min	127.8553390572779
空間範圍.X.max	127.88097380893306
空間範圍.Y.min	26.630362980584657
空間範圍.Y.max	26.68047930832328

### 管理資訊


產製者	Tzu-Hao Lin, Tomonari Akamatsu, Frederic Sinniger, Saki Harii
聯絡人	Tzu-Hao Lin






 **Social**


---

 Twitter

---

 Facebook


---

 **License**


---

[CC-BY 4.0](#) [OPEN DATA](#)

---

 **Cite as** Beta


---

American Psych... 

---

Tzu-Hao Lin, Tomonari Akamatsu, Frederic Sinniger, Saki Harii. (2021). *Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan* (Version 2021-01-09T09:11:31.023608) [Data set]. Retrieved from <https://data.depositar.io/en/dataset/coral-reef-sesoko>

---

 **Cut to clipboard**

Deployment and recovery of recorders were conducted by divers.

**Data Processing**





Audio recordings generated by AUSOMS-mini recorders were saved in MP3 format. Each MP3 is about 8-hour long and do not have a time stamp on the file name. To facilitate data management, we segmented the 8-hour long MP3 into WAV files of 5-min duration.


We used the [LTSA\\_gui](#) to generate long-term spectrograms (LTS) and save the LTS in mat files. Each mat file contains median-based LTS and mean-based LTS. Median-based LTS was obtained by measuring median power spectral densities within each 5-min segment. Mean-based LTS was obtained by measuring mean power spectral densities within each 5-min segment.

**Associated Publication**

Tzu-Hao Lin, Tomonari Akamatsu, Frederic Sinniger, Saki Harii (2020) Exploring coral reef biodiversity via underwater soundscapes. [Biological Conservation](#), 253: 108901.

**Data and Resources**

	<b><a href="#">Audio data</a></b>	<a href="#">Explore</a>
A link to a shared Drive folder of underwater recordings (WAV) and long-term...		
	<b><a href="#">Long-term spectrogram of Site A</a></b>	<a href="#">Explore</a>
A mat file contains the median- and mean-based long-term spectrograms.		
	<b><a href="#">Long-term spectrogram of Site B</a></b>	<a href="#">Explore</a>
A mat file contains the median- and mean-based long-term spectrograms.		
	<b><a href="#">Long-term spectrogram of Site C</a></b>	<a href="#">Explore</a>
A mat file contains the median- and mean-based long-term spectrograms.		







Contents lists available at ScienceDirect

Biological Conservation

journal homepage: [www.elsevier.com/locate/biocon](http://www.elsevier.com/locate/biocon)

## Exploring coral reef biodiversity via underwater soundscapes

Tzu-Hao Lin<sup>a,\*</sup>, Tomonari Akamatsu<sup>b,\*\*</sup>, Frederic Sinniger<sup>c</sup>, Saki Harii<sup>c</sup><sup>a</sup> Biodiversity Research Center, Academia Sinica, Taiwan<sup>b</sup> The Ocean Policy Research Institute, The Sasakawa Peace Foundation, Japan<sup>c</sup> Tropical Biosphere Research Center, University of Ryukyus, Japan

## ARTICLE INFO

## Keywords:

Ocean sound  
Mesophotic corals  
Remote sensing  
Noise  
Acoustic habitat  
Acoustic diversity

## ABSTRACT

Information on biodiversity is essential to evaluate the ecological status of coral reefs. Sounds produced by reef-associated organisms have been used as a biodiversity indicator. However, the interference from abiotic sounds and the lack of a comprehensive audio library have impeded effective evaluation. This study investigated the application of underwater soundscapes as a remote-sensing method to detect biological and anthropogenic activities. Using techniques including the visualization of long-duration recordings, source separation, and clustering, soundscapes were separated into sounds of anthropogenic and biological sources. Our results revealed the dynamics of biological sounds among coral reefs off Sesoko Island, Oki nawa, Japan. Biological sounds were much more prominent in shallow-water reefs than in upper-mesophotic reefs, but their spectral features and compositions differed. The shallow-water reefs were dominated by broadband sounds of crustaceans and low-frequency transient fish calls, whereas the upper-mesophotic reefs were characterized by a diverse array of fish choruses and transient sounds. We also discovered that shipping noise heavily interfered with the soundscapes from the upper-mesophotic reefs and represented an invisible threat to life in the low-light habitat. The applied techniques of soundscape information retrieval revealed the distinct ecological status of coral reefs and the behavior change of sound-producing organisms in high temporal resolution. Implementation of soundscape monitoring can generate ecological information on habitat quality, reef biodiversity, human activities, and their interactions. Global collaboration on underwater soundscapes will establish a data-informed platform and help stakeholders assess the resilience of coral reefs to environmental and anthropogenic stressors.

## 1. Introduction

Marine ecosystems provide irreplaceable services and currently face significant pressures due to climate change, human disturbance, and excessive use of marine resources. The United Nations has recognized these threats and placed the conservation of marine ecosystems as one of its sustainable development goals (UN General Assembly, 2015). Coral reefs support various social and economic activities, such as fisheries, coastal protection, and tourism, of many maritime tropical and subtropical nations (Moberg and Folke, 1999; Barbier, 2017; Spalding et al., 2017; Woodhead et al., 2019). These benefits rely on the abundant biodiversity in coral reefs. However, coral reefs have undergone recurrent high-frequency bleaching episodes over the past 20 years due to increased sea surface temperatures (Hughes et al., 2017, 2018). Therefore, detailed information on the spatiotemporal changing patterns of marine biodiversity and interactions with human activities is crucial for

the conservation management of coral reefs.

Biodiversity monitoring in coral reefs remains challenging, partially due to the distinct reef environments and their unique fish assemblages (Pearman et al., 2018; Dumalagan et al., 2019). A comprehensive and long-term assessment of reef biodiversity, environmental characteristics, and human activities may not be feasible because of limited resources for observation and survey opportunities, especially for developing regions or remote reefs. An underwater sensing system capable of monitoring the changing patterns of marine biodiversity, with the ability to diagnose potential risks due to environmental and anthropogenic stressors, is required for establishing management strategies of coral reefs and for providing alerts to the early-warning signs of ecosystem changes (Schmeller et al., 2017; Obura et al., 2019).

A potential solution for such an underwater sensing platform is through monitoring ocean sounds. One autonomous recorder can store long-duration audio recordings, with improved time resolution of

\* Correspondence to: T.-H. Lin, Biodiversity Research Center, Academia Sinica, 128 Academia Road, Sec. 2, Nankang, Taipei 11529, Taiwan.

\*\* Correspondence to: T. Akamatsu, The Ocean Policy Research Institute, The Sasakawa Peace Foundation, 1-15-16 Toranomon, Minato, Tokyo 105-8524, Japan.  
E-mail addresses: [lintzhuo@gate.sinica.edu.tw](mailto:lintzhuo@gate.sinica.edu.tw) (T.-H. Lin), [akamatsu.tom@gmail.com](mailto:akamatsu.tom@gmail.com) (T. Akamatsu).

<https://doi.org/10.1016/j.biocon.2020.108901>

Received 18 June 2020; Received in revised form 23 November 2020; Accepted 27 November 2020

Available online 10 December 2020

0006-3207/© 2020 Elsevier Ltd. All rights reserved.

in mesophotic  
s. However, the  
n shallow-water  
ound-producing  
Although basic  
remains scant,  
ill improve our  
lerstudied deep-  
s multiple mes-  
ers to prioritize

asurement of the  
n shallow-water  
horus may be a  
which has been  
may be partly  
the c  
n et  
water  
g wir  
be eas  
t may

system  
er so  
larva  
d set  
(Ver  
the r  
hance  
erfere  
d the  
the re  
Moon

and n  
to sh  
eries o  
photi  
of a b  
prop  
ctivit  
esoph  
er the  
sonit

d ocean temper-  
teriorate due to  
on. Therefore, it  
silience of coral  
nagement plans  
ed here and the  
ystem functions

fs

t from the long-  
ployment of an  
rk methods, and  
als. This study  
ous acoustic reef  
ef soundscapes.

With the recent development of underwater technology and audio information retrieval techniques, a soundscape monitoring network can generate numerous acoustic data that contain ecological information in multiple dimensions, including the quality of the acoustic habitat, community of sound-producing organisms, and potential effects due to human activities. The generated information will allow managers and stakeholders to conduct a more comprehensive assessment of ecosystem health at scale.

The feasibility of using 2-h daytime recordings to obtain rapid ecoacoustic indicators to characterize the major ecosystem functions of coral reefs is currently being discussed (Elise et al., 2019a). Our result suggests that beyond a spatial survey based on snapshot recordings, several long-term recording sites must be established within a spatial range of interest to allow a full understanding of the temporal variations of underwater soundscapes and ecosystem dynamics. The exceptional power of time-series recordings can help establish a library to document

## Data availability

The audio dataset used in preparing this paper are available from the corresponding authors on reasonable request. A dataset of the LTS is available on depositar (<https://data.depositar.io/en/dataset/coral-reef-sesoko>).

### Data availability

The audio dataset used in preparing this paper are available from the corresponding authors on reasonable request. A dataset of the LTS is available on depositar (<https://data.depositar.io/en/dataset/coral-reef-sesoko>).

### CRediT authorship contribution statement

**Tzu-Hao Lin:** Conceptualization, Methodology, Software, Validation, Data curation, Formal analysis, Resources, Writing – original draft.  
**Tomonari Akamatsu:** Conceptualization, Methodology, Resources, Data curation, Writing – reviewing and editing, Funding acquisition.  
**Frederic Sinniger:** Conceptualization, Visualization, Investigation, Data curation, Writing – reviewing and editing.  
**Saki Harii:** Conceptualization, Investigation, Writing – reviewing and editing, Funding



## Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan

Followers




0

### Project



### Ocean Biodiversity Listening Project

Project Website The ocean is full of sounds that are generated from geophysical events, marine animals, and human activities. By using a hydrophone (a microphone for underwater... [read more](#)

Dataset  Topics  Activity Stream  History

## Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan



This dataset is an archive of audio data of shallow-water and upper-mesophotic coral reefs off Sesoko Island, Okinawa, Japan. Python codes to visualize the audio data were also provided in a notebook based on Google Colab.

### Recording Locations

Three long-term recording sites were established since May 2017. Site A (N26.635° E127.865° ) is located on the southeast coast of Sesoko Island and in front of the Sesoko Station of the University of the Ryukyus. The water depth is 1.5 m. Site B (N26.665° E127.869° ) is located at the bottom of a reef slope on the north of Sesoko Island and the west of Toguchi Port. The water depth is 20 m. Site C (N26.670° E127.866° ) is located on a nearly flat plateau to the north of Sesoko Island and the west of Toguchi Port. The water depth is 40 m.

### Acoustic Recorders

[AUSOMS-mini stereo recorders](#) (AquaSound, Kobe, Japan) were used to collect underwater sounds. From May 2017 to July 2018, six AUSOMS-mini recorders were used: 14-0106, 14-0107, 15-0106, 15-0107, 15-0109, 15-0110.

### Configuration of Audio Recording

(1) Duty Cycle: continuous. (2) Sampling Rate: 44.1kHz. (3) Channels: 2. (4) File Format: MP3 (128 kbps). (5) Audio Gain: High. (6) High Pass Filter: Off.

### Field Deployment

At each recording site, one AUSOMS-mini stereo recorder was fixed to a cement

# Data Discovery via Google Dataset Search

<https://datasetsearch.research.google.com/search?query=Coral Reef Soundscapes>

The screenshot shows a Mozilla Firefox browser window displaying the Google Dataset Search results for the query "Coral Reef Soundscapes". The browser's address bar shows the URL: <https://datasetsearch.research.google.com/search?query=Coral Reef Soundscapes&docid=L2cvM>. The search bar contains the text "Coral Reef Soundscapes".

At the top of the results page, there are filters for "Last updated", "Download format", "Usage rights", "Topic", and "Free". A "Saved datasets" button is also visible.

On the left side, a sidebar indicates "29 datasets found". The first result is "Data from: Hurricane impacts on a coral reef soundscape" by NC State University, with a "txt, zip" format and an update date of Dec 28, 2020. The second result is "Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan" by data.depositar.io, with a "mat" format and an update date of Jan 9, 2021. The third result is "Correlation between benthic algal cover and coral reef soundscapes" by PLOS.

The main content area displays the details for the selected dataset: "Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan". A blue button labeled "Explore at depositar" is highlighted with a yellow arrow. Below this button, a list of dataset identifiers is shown: `mat(151517946), mat(178270495), mat(141770285)`. The dataset was updated on Jan 9, 2021, and is provided by the Ocean Biodiversity Listening Project. The license is Attribution 4.0 (CC BY 4.0). The description states: "This dataset is an archive of audio data of shallow-water and upper-mesophotic coral reefs off Sesoko Island, Okinawa, Japan. Python codes to visualize the audio data were also provided in a notebook based on Google Colab." The recording locations are described as three long-term sites established since May 2017: Site A (N26.635° E127.865°), Site B (N26.665° E127.869°), and Site C (N26.670° E127.866°).

## Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan

Followers



0

### Project



### Ocean Biodiversity Listening Project

Project Website The ocean is full of sounds that are generated from geophysical events, marine animals, and human activities. By using a hydrophone (a microphone for underwater... [read more](#)

Dataset  Topics  Activity Stream  History

## Coral Reef Soundscapes off Sesoko Island, Okinawa, Japan



This dataset is an archive of audio data of shallow-water and upper-mesophotic coral reefs off Sesoko Island, Okinawa, Japan. Python codes to visualize the audio data were also provided in a notebook based on Google Colab.

### Recording Locations

Three long-term recording sites were established since May 2017. Site A (N26.635° E127.865° ) is located on the southeast coast of Sesoko Island and in front of the Sesoko Station of the University of the Ryukyus. The water depth is 1.5 m. Site B (N26.665° E127.869° ) is located at the bottom of a reef slope on the north of Sesoko Island and the west of Toguchi Port. The water depth is 20 m. Site C (N26.670° E127.866° ) is located on a nearly flat plateau to the north of Sesoko Island and the west of Toguchi Port. The water depth is 40 m.

### Acoustic Recorders

[AUSOMS-mini stereo recorders](#) (AquaSound, Kobe, Japan) were used to collect underwater sounds. From May 2017 to July 2018, six AUSOMS-mini recorders were used: 14-0106, 14-0107, 15-0106, 15-0107, 15-0109, 15-0110.

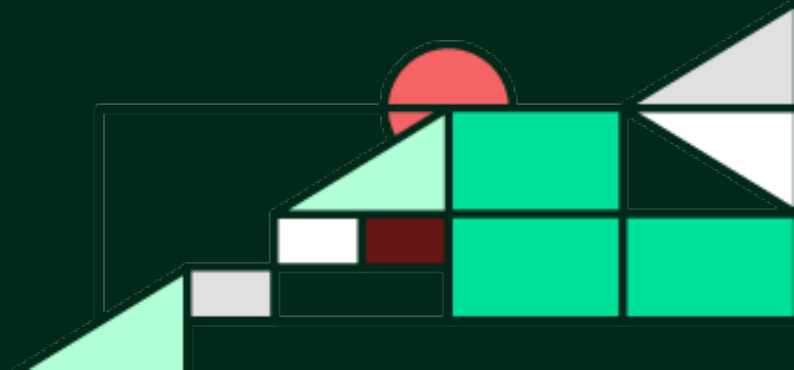
### Configuration of Audio Recording

(1) Duty Cycle: continuous. (2) Sampling Rate: 44.1kHz. (3) Channels: 2. (4) File Format: MP3 (128 kbps). (5) Audio Gain: High. (6) High Pass Filter: Off.

### Field Deployment

At each recording site, one AUSOMS-mini stereo recorder was fixed to a cement

# A bottom-up data repository



# The road to *depositar*

- Developed to facilitate data sharing for two “integral” research projects funded by the Ministry of Science and Technology (MOST), Taiwan
  - 2013 – 2015 and 2016 – 2019
  - Presented at Open Repositories 2015: *Building Research Data Repositories for the Humanities and Area Studies with CKAN (and Some Extensions)* – Chuang, Lee, Syu, and Yeh
- Open to the public for general use in Oct 2018
  - Launched at the Pacific Neighborhood Consortium 2018 Annual Conference (co-located with Digital Heritage 2018), San Francisco, CA, USA
  - Change of role: data publisher for some → data depository for all



# Organic bottom-up growth

- Uptake by NPOs (non-profit organizations) since 2019
  - e.g. from the groups monitoring ecological impacts resulted from governmental construction works
- New grant from MOST with some focus on Research Data Management (2019 – 2022)
  - interacting with other project teams working on Sustainability Research and on Long-Term Social and Ecological Research (i.e. “LTSER” stations)
- Some use by researchers at Academia Sinica, Taiwan
  - as well as others in Taiwan and from abroad
- *Terms of Use and Privacy Policy*, July 2021



## Workshops co-organized by NGOs on Research Data Management (RDM)

# 2019 公共工程生態檢核 資料管理與開放應用工作坊

2019年11月02日：台北南港中央研究院人文社會科學大樓遠距會議室(台北市南港區研究院路二段128號)

2019年11月09日：台中文山社區大學—台中世貿中心3樓303會議室(台中市西屯區天保街60號)

2019年11月23日：台南社區大學—後甲國中校區「池東樓112教室」(台南市東區林森路二段260號)

議程 線上報名 範例資料集

2017年4月25日公共工程委員會頒佈公共工程生態檢核機制；2019年5月10修正，重新頒佈為公共工程生態檢核注意事項，規範中央目的事業主管機關的新建工程，或地方政府接受中央經費補助超過50%的新建工程，從計畫核定至工程完工階段必須執行生態檢核，並將相關資料、資訊公開於開放平台。2017年開始的前瞻建設水環境計畫也必須依此規範執行生態檢核，落實資訊公開和公民參與。目前，各單位仍缺乏妥善管理、開放生態檢核相關資料的協作平台，以致前瞻水環境建設計畫的資訊公開不完整、不透明，執行單位亦忽視生態檢核程序，造成許多水環境改善、營造計畫破壞既有自然環境而引起公眾嚴厲的批評。因此，公民團體要求相關單位必須立即改善資訊公開作為，開放生態檢核相關資料。

研究資料寄存所是一個通用的研究資料管理協作平台。此平台短期內可以協助提供政府相關部門管理和開放生態檢核資料所需的平台與技術資源；長期維運，更可以持續妥善保存全國的生態環境基礎資料，提供未來各領域學術研究、政府政策檢討及民間加值運用，是全民共享的重要數位資產。目前，研究資料寄存所已累積至少106個與生態檢核相關的資料集；但是，多數資料集缺乏基本的詮釋資料，且上傳的資料多為PDF檔，可用性低，對於公民團體要求部門落實資訊公開和資料開放的期許仍有很大的落差。為提升資料提供者(相關政府單位、水環境顧問團及執行生態檢核廠商)對於資料管理的觀念與使用研究資料寄存所妥善管理資料的能力，並向一般民眾和公民團體推廣使用此平台來瞭解公共工程對環境生態的影響，我們邀請中央目的事業主管機關、縣市政府、前瞻水環境計畫顧問團、生態檢核計畫團隊、生態顧問公司及民眾與公民團體，就近參加11月2、9、23日分別在台北、台中、台南舉辦的三場生態檢核資料管理與應用工作坊，以提升生態檢核資料管理的品質，並促進公共工程資訊公開與公民參與。

中央研究院資訊科學研究所、中央研究院資訊科技創新研究中心、中央研究院網格計算中心、中央研究院地理資訊科學研究專題中心、台中文山社區大學、台南社區大學、台灣生態學會、竹圍工作室、台灣河溪網、台灣石虎保育協會 合作舉辦

## Translation of the RDM Guide from Science Europe

### 國際合用的研究資料管理實用指南 增訂版

Practical Guide to The International Alignment of Research Data Management - Extended Edition

包含 DMP 評量指標  
with DMP Evaluation Rubric

Science Europe



# Building on top of CKAN

- CKAN from OKFN: open source and in Python
  - repurposed and customized for research datasets
  - enriched metadata & connected to more resources

- Added features

- Wikidata (2017)



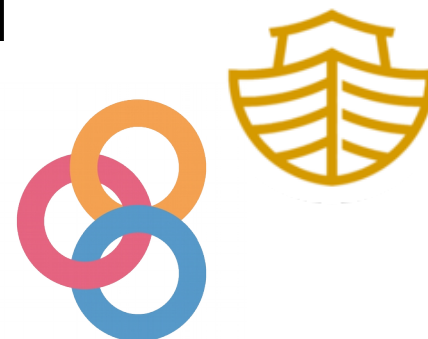
- Citation Widget (2018)



- RDF Serialization (2020)

- Archival Resource Key (2022 Q2)

- BinderHub Integration (2022 Q2)





# Discussion

- Paths to sustainability
  - Tools (keep no data) →  
Systems (keep our data) →  
Services (keep other people's data)
  - Persistent services and/or persistent datasets?
- Collaborations among diverse data communities
- Pooling resources together

# Thank You!

<https://data.depositor.io/>

[data.contact@depositor.io](mailto:data.contact@depositor.io)

The *depositor* is a collaboration at the Institute of Information Science, the Research Center for Information Technology Innovation, and the Research Center for Humanities and Social Sciences (GIS Center) in Academia Sinica, Taiwan. The project has been supported, in part, by grants from Taiwan's Ministry of Science and Technology.

「研究資料寄存所」是中央研究院資訊科學研究所、資訊科技創新研究中心、人文社會科學研究中心（地理資訊科學研究專題中心）的協作專案，部份經費來自台灣科技部的專題研究計畫。

