SEARPP Data Policy

With effect from the 1st January 2024, we will require all researchers working with SEARRP to *deposit the most recent version* of all datasets generated and collected as part of the programme. All datasets provided must have accompanying *completed metadata* to enable and simplify interpretation. All deposited datasets will be stored on Zenodo, which enables open access to the research data and provides a permanent DOI for each dataset. The full list of data submitted can be seen on our Zenodo community page:

https://zenodo.org/communities/searrp/

We have three reasons for holding data in a centralised data repository:

1. The data depository ensures we retain all information relevant to the current research projects being implemented with SEARRP. Over time, as new researchers work with SEARRP and deposit their datasets, historical data records will become progressively more valuable as a means of estimating ecological changes through time. Without a central database, these records would become progressively more fragmented and distributed among research groups around the world, making it virtually impossible to track down data in the future.

2. We expect researchers will benefit directly from the database. It provides an opportunity for researchers to obtain data to support their analyses without having to collect it themselves.

3. Furthermore, having a centralised database, we provide a platform for researchers to link data collected from their research project with other datasets available in the depository to explore broader research questions. This also provides a means to support collaborations and interdisciplinary projects between researchers.

Depositing data

Researchers are required to deposit their respective datasets and metadata *within twelve months of data collection*, or *immediately post-publication for datasets that get published*. For multi-year projects, we expect researchers to *provide annual updates* to the central database.

To submit, please send your datasets via email to our Data Manager, Sam Lee at samleext@gmail.com, and he will upload the datasets into the data validator. In the event that your dataset contains errors, an error report will be sent to you via email and you will need to revise the submission before it is publishable. For formatting instructions, please refer to the Submission Process wiki on the SEARPP website.

We currently only support data submitted as Excel (.xlsx) files. If you have data in other formats or any general enquiries, please contact Samuel Lee at samleext@gmail.com to ask for guidance.

Data protection

The metadata for submitted data will be publicly accessible through Zenodo.

If data is submitted as open access, it will be made publicly available from Zenodo automatically. The DOI and publication details of each dataset provide a permanent citation for the dataset, that users of your data are expected to cite. We also use the existing Zenodo mechanism for publishing embargoed data. When you submit a dataset, you can embargo that data for up to two years. The embargoed data is stored on Zenodo and the metadata is made available, but the data will only become freely available once the embargo date has passed.

If you use any open access SEARPP data, please cite the relevant researchers who collected the original dataset (or the associated publication if that is more appropriate) and consider including the PIs of the data as co-authors on manuscripts. The inclusion of data collectors as co-authors is obligatory for long-term datasets that are field intensive (e.g., SBE data, 50-ha ForestGEO and animal camera trapping data). If you would like access to an embargoed dataset, requests should initially be directed to the individual researchers who collected the particular data you wish to use.

In certain cases, specifically long-term intensive datasets, the existing Zenodo mechanism can be used for sharing data to protect the interests of both data collectors and data users. When you submit a dataset, you can define the terms and conditions for its release to a third party. However, this pathway should be limited to continually collected data and not projects that were short-term and have reached completion.

Current & historic datasets

Over time, and in collaboration with previous SEARRP scientists, we plan to compile historic datasets collected prior to the implementation date (1st January 2024) of this new policy. We can also link previously stored data in other repositories to our Zenodo page by providing a basic description and DOI link to the stored data. This process will not generate a new DOI, but aid in providing a reference on the SEARRP Zenodo page to previous datasets. Michael O'Brien and Samuel Lee will be contacting SEARRP scientists in the coming months to discuss this in further detail.