Ecological Impact Assessment - FAIR Principles Compliance

**Findable**

Each dataset included in this project has been assigned a globally unique identifier. Each sampling event has been assigned a unique event ID according to the location and method of the sample. Each species occurrence has been assigned a unique occurrence ID according to the event ID location. Species have been labelled clearly. Explanations of the data, including how it was generated and ID assignments have been clearly described by metadata in the excel datasheet. The identifier of each unique datasheet has been included in the metadata to explicitly link the descriptions to the data it is describing. The datasets can be easily searched within the excel spreadsheet according to these unique identifiers or based on specific criteria such as species. The data can be easily found and referenced.

**Accessible**

All data and metadata obtained and created during this project is publicly available on GitHub. The data is open, free to access and universally available. Data is fully accessible to be downloaded and authenticated without authorisation procedure. Data is not restricted. Metadata are available within the dataset for findability, and is available in a separate file on GitHub, so metadata is accessible even if the data was no longer made available in order to understand the data and the context in which it was generated.

**Interoperable**

Language used in this dataset and metadata are standardised and formal to allow integration with other resources or datasets. Datasets are stored in excel and CSV formats, which are widely supported across different softwares and can be easily shared. Analysis is stored in an R script which is broadly used across the field. Language used is consistent across the dataset and metadata and follows the same standards as GBIF template. Vocabularies used are consistent and follow FAIR principles. Terminology in datasets and metadata can be easily searched to find other datasets and is freely accessible.

**Reusable**

Data is stored in a standardised format according to the GBIF data template and GBIF taxonomy allowing data to be reused for different projects, for data to be combined with different datasets and to be manipulated using different software and sources. Datasets have been formatted using Darwin Core terminology and terminology to aid data reusability. Event IDs have been standardised across the dataset according to location and sampling site. Location names have been standardised and international standard alpha-2 country codes have been used. Collected data is described in detail, aiding reusage and integration into different sources.