Automated classification of Hainan Gibbon call types using deep learning - Student Outline DMP

---

1. General guidelines

PURPOSE OF THIS TEMPLATE - The purpose of the Outline DMP is to indicate your initial plans for how your data will be collected, shared and stored, and to give you a chance to think about these data-focused aspects of the research process. As your begin doing your research, your data process may change, and it is perfectly acceptable to change your data management plan to accommodate the changes in your research process. Indicate below that you understand the purpose of completing this Outline DMP template.

I understand the Outline DMP template is a projection of my anticipated data management planning requirements and should be updated as my project develops.

2. Authors and supervisors

PROJECT NAME - Replicate the title of your project, dissertation or thesis exactly as it appears in your proposal document.

Automated classification of Hainan Gibbon call types using deep learning.

PERSONAL DETAILS - Indicate the name(s) and student number(s) of the student(s) who will be involved in this project, dissertation or thesis.

Name: Nonhlanhla Luphade
Student Number: LPHNON003

SUPERVISOR(S) DETAILS - Indicate who will supervise this project, dissertation or thesis. If you do not yet have a supervisor, leave this section blank.

Supervisor: Prof. Ian Durbach
Co-Supervisor: Steffan Britz

3. Data Collection/Generation

COLLECTION OF ORIGINAL DATA - Indicate whether or not you intend to gather/produce original data for your study, and provide a brief description of the kind of data you think you will collect. If you are unsure at this time, indicate what you think you are most likely to collect. If you are not intending to gather or collect your own data, declare that here.

I intend to use publicly available data.

USE OF EXISTING DATA - Indicate if you intend to re-use existing data, either from online searches or from datasets provided by your supervisor, lab, or funder. If you are not intending to re-use existing data, declare that here.

- Existing data will be used for the research, subsets of the data are stored on Zenodo.
- The rest of the data will be provided by the supervisor.

- Permission to use the data has been granted by the supervisor.

DATA SHARING - Indicate whether or not you are intending to publish your research data. If you are, indicate where you are intending to publish your data and under what licensing conditions, such as Creative Commons. If you are not intending to publish your data, provide reasons and reference the appropriate ethical considerations, commercial applications/patenting ambition, or data re-use agreements that prevent you from publishing your data.

The data is publicly available.

4. Data Storage

ANTICIPATED DATASET SIZE - Indicate the estimated size of your completed dataset, and indicate whether or not you will need to access additional data storage facilities. If such storage is not provided by your unit or department, you may need to factor in the cost of purchasing additional storage space.

- Between 20GB and 100GB

The data is stored on the cloud (google drive, one drive, etc.).

DATA BACKUPS - Indicate how you plan to ensure your data is secure and retrievable in case of errors or hardware failure. Describe what procedures you will put in place to back-up copies of your data and where they will be stored.

Not applicable

5. Data Centre(s)/Repositories

DATA CENTRES/REPOSITORIES - Once your project, dissertation or thesis is complete, it is advisable to curate and archive your completed dataset with an established data centre or repository. Check with your supervisor or funder if you are required to deposit your data in a specific repository, or declare that you will deposit the data in ZivaHub (see the Guidance section).

The links to the data and the output from the thesis such as predictions, generated models, etc will be stored on the UCT repository ZivaHub at the end of the thesis.

METADATA - Metadata is descriptive information that others will need to make sense of your dataset. Metadata includes things like study descriptions or abstracts, study instruments (sample collection schedules, codebooks for variables, survey instruments, etc.), subject codes, and keywords. Indicate what metadata will accompany your curated dataset.

The completed dataset will be accompanied by keywords, a short description taken from my dissertation abstract, and relevant paragraphs on the data process taken from my methods section.

6. Budget

BUDGET - Indicate any costs specifically relating to the management and curation of your data, such as purchasing additional storage space, digitisation of physical media, data storage or curation charges, and data audits. Most student research will be able to make use of free options provided by UCT and will not have to budget for data costs.

Not applicable.

Created using the DMPRoadmap. Last modified 09-03-2021

3 of 3