**Data Management Plan for Professor Periwinkle**

Data Needing to be Created or Collected

Licensing Information

Necessary Facilities/Equipment

Recommended Data Management Practices

Ownership and Access to Data

Post-Project Data Value

Metadata Requirements

Relevant Linked Open Data Strategies

Reuse and Long-Term Preservation

Storage Costs

**Dear professor Periwinkle,**

Below you will find a detailed data management report identifying the resources required for your upcoming project on tracking marine wildlife. We hope that this plan serves you and your team well and meets all your data needs.

Sincerely,

Xx

**Purpose**

The purpose of this data management plan is to make recommendations to Dr. Periwinkle on how best to achieve her long-term goal of making data from her lab available for use outside the university. Specifically, the data collected is used to write papers and to create exercises for graduate classes. However, the is currently no structured sharing of data (other than visually) from her lab outside of the university. In the future, Dr. Periwinkles goal is to make this data more accessible and available for use.

**What data will be created or collected (type, size, format, etc.)**

The data collected will be that of marine wildlife to help write papers and create exercises for graduate level classes. Professor Periwinkle’s team will be using a variety of sensors and monitoring equipment to collect the data, including: i) digitally by remotely-operated marine vehicles (ROMV), ii) through tags that are surgically implanted in captured and released animals, iii) by static sensor buoys that measure ocean conditions, and iv) through communication lines that passively listen for signals for animal tags. Additional forms of data that are currently collected include field notes that document the animals captured and tagged, and field notes from observational studies conducted during classes Dr. Periwinkle teaches. Currently, Dr. Periwinkle’s website also allows citizen scientists to report sightings of wildlife in the Minas Basin, although this data is not currently used.

In terms of the size and format, each collector produces data in a format specific to the sensor, however, Dr. Periwinkle’s team has developed software that converts the data to the NetCDF format for the use and storage. NetCDF is a set of software libraries and self-describing, machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data (UNIDATA). The sensors are currently producing roughly 300 MB per day in raw sensor data, which becomes 500 MB in uncompressed NETCDF formatted data. It is important to note that both versions are kept. Additionally, Dr. Periwinkle will be looking to use the data from her website in the future, which is downloaded as .tsv files and is 500GB. Additionally, members of Dr. Periwinkle’s team run complex simulation models that attempt to predict animal populations and movement based on oceanographic data acquired from the buoys and from collaborators. These models produce roughly 200GB of data in zipped .csv files.

Also has storage from 3 years on floppy disks. Add on the 500 GB of website data and 200 CB of csv data

At a baseline, Dr. Periwinkle

**• What licenses apply to the data**

As stated, the data collected will be used to write papers and to create exercises for graduate classes. The data will be made available in visual form to the general public using OceanView.org, a platform for sharing oceans-related data. However, there is currently no structure sharing of data (other than visually) outside of the university. The intent and long-term goal is to make the actual data available for use. Additionally, there is currently no specific licensing in place for students. The citizen scientist’s data is open and there is no procedure or contract in place that the students (incoming or leaving) have to sign. In contrast, Dr. Periwinkle owns the data she produces.

With this, it is recommended that Dr. Periwinkle use Creative Commons Attribution required (CC BY) licensing through a public domain in order to allow others to freely access her papers, and other data, without having to request the data files.

**• What facilities and equipment will be required (hard disk space, backup**

**server, central repository, off-site repository, etc.)**

Currently, data in Dr. Periwinkles lab is primarily shared with the team using external hard drives and USB keys**.**

**• What data management practices (backups, storage, access control,**

**archiving etc.) will be used**

Currently, past data and current data is kept in Dr. Periwinkles office in a myriad of forms including: floppy disks, ZIP Disks, CDs, DVDs, Blu-Ray, external hard drives, as well as through Dropbox.

Backups and archiving she would need a physical hard drive. (in addition to cloud data). If they lose their licenses, their cloud data could be deleted, so this could be an additional security measure.

**• Who will own and have access to the data**

Currently, Dr. Periwinkle is the owner of her own data whereas the data produced by the citizen scientists is open.

**• Which data will retain value after the life of the project**

The field notes follow the Darwin Core: <http://www.iobis.org/manual/darwincore/>  
All the required filed must be filled, otherwise the data are dismissed.

**• What metadata and linked open data strategies will be employed**

**• How will its reuse be enabled and long-term preservation ensured after**

**the original research is completed**

**• How much will the storage of this data cost (cloud and/or hard drives)**

Dr. Periwinkle has funding in her grant for the Data Management Plan.

<https://www.unidata.ucar.edu/software/netcdf/>

**Our Recommendation**

As stated, the data collected is used to write papers and to create exercises for graduate classes, however, there is currently no structured sharing of data outside of the university. The long-term goal is to make this data available for use. In order to do so, in conjunction with the information from this data management plan, we recommend that Dr. Periwinkle …

**The problem**

Someone looking to access Dr. Periwinkles work can email Dr. Periwinkle to request the data files, and Dr. Periwinkle can upload them to Dropbox. It is important to note that although the university uses the file sharing provider Microsoft OneDrive, it requires external users to login with a Microsoft-provided account of some kind and not all people have these accounts.

**Professor Pinkerton**

Necessary Data

Licensing Information

Necessary Facilities/Equipment

Recommended Data Management Practices

Ownership and Access to Data

Post-Project Data Value

Metadata Requirements

Relevant Linked Open Data Strategies

Reuse and Long-Term Preservation

Storage Costs

**Professor Chartreuse**

Necessary Data

Licensing Information

Necessary Facilities/Equipment

Recommended Data Management Practices

Ownership and Access to Data

Post-Project Data Value

Metadata Requirements

Relevant Linked Open Data Strategies

Reuse and Long-Term Preservation

Storage Costs