Plant Herbivory by Species in Guam

1. RDM Documentation Table

This table is to be filled out as completely as possible before the beginning of the project, and updated as needed, including at the end of the of the project, and after, as derivative publications are created.

2. Description of the Data

- What type of data will be produced? (Tabular, sensor, video, audio, etc.)
- In what formats? (.txt, .cvs, .tiff, etc.)
- Are there special tools or software needed to create/process/visualize the data?

A collection of .jpeg and .tiff files will be collected of leaf scans. The area will then be determined using PowerPoint as well as the area of the leaf eaten. These numerical values will then be inputed into a .csv file. R will then be used to conduct a statistical analysis on the data collected.

- How will data be collected?
- How will the data collection be documented?
- What project and data identifiers will be assigned?
- Will you use pre-existing data? If so, from where?

Data will be collected by using SCION to scan leaves and then stored in a computer as .jpg files. After determining the area of the leaves and the percent herbivory, the data will be stored in a .csv file.

- How much data will there be, and at what growth rate? (1 GB, 3 TB etc.) How often will it change?
- How will you store, backup, and protect data from lost during the research project?
- Who is responsible for managing the data?

The data is about 30 MB and will not change. The data is backedup on GitHub and Marjorie Hanneman and Haldre Rogers are responsible for managing the data.

- Will the data collected be unique or will the data be reproducible? What would happen if the data got lost or became unusable later?
- Who will potentially use the data?
- What value does the data have over the long- term? (Please consider not only your research team, but third parties as well.)

The data collected will be unique and if the data is lost the project will start over.

3: Standards to be Used

- Are your data formats open or proprietary? If proprietary, what is the rationale for using that format?

The data formats are open and will be in .jpg, .tif, and .csv. These will allow easy analysis in R and will be easy to use in the future.

- What standards will be used for documentation and metadata?
- What documentation or descriptive metadata will you be creating in order to contextualize the data for future users?

4. Data Organization and Description

- How will the data be organized?
- What directory and file naming conventions will be used?

The data will be organized in tables that are made tidy using R. They will be under a repository on GitHub named MarjorieHan.

- What metadata schemas are appropriate for describing these types of data?
- What metadata schema will be chosen for this data?

5. Policies for Access

- Does the data contain any personally identifiable information (PII)?
- If so, how will you anonymize or deidentify the data if PII is present?

This data does not contain any personally identifiable information.

- Are there any special privacy or security requirements (e.g., personal data, high-security data)?
- What steps will be taken to protect privacy, security, confidentiality, intellectual property or other rights?
- Does your data have any access concerns? Describe the process someone would take to access your data.

There are no special privacy or security requirements for this data. To protect intellectual property, the data is only shared with those working on it.

Who controls the data (e.g., funder, PI, student, lab, University)?

Are there any embargo periods to uphold?

Haldre Rogers and affiliated individuals control the data.

6. Policies for Re-Use, Re-Distribution, and Derivative Products

Who might be the audience for data reuse? Who will use the data now? Later?

Any sharing requirements (e.g., funder data sharing policy)?

When will the data be published and where?

If you allow others to reuse your data, how will the data be discovered and shared?

What license type is being used (e.g.: Creative Commons 0, etc.)

What special tools and/or software are needed to work with data?

7: Plans for Archiving and Preservation

Are there data archives that my data is appropriate for (subject-based? Or institutional)?

Which archive will the data be stored in and why was it chosen?

What is the persistent identifier type used by the archive?

How will the data be archived for preservation and long-term access?

How long should it be retained (e.g., 3-5 years, 10-20 years, permanently)?

What file formats will the data be preserved in? Are they long-lived? Are they proprietary?

Who will maintain my data for the long-term?