**EXAMPLE DATA MANAGEMENT PLAN: BIOLOGY**

**1. Research Products, Data Format and Data Sharing:** This project will generate research products on climate variability, secondary metabolite chemistry of tea leaves from *in situ* stratified sampling and manipulative experiments, and from farmer and consumer surveys. The specific products, format of data storage, and format of data sharing is listed below. All the project data will be stored using Tufts University Information Technology (UIT) resources. A main study deliverable is a web-accessible catalogue of all data products generated from this project posted on a Spark 2.0 wiki infrastructure maintained by Tufts UIT. UIT offers a suite of Web 2.0 tools for communication and collaboration across Tufts University entitled Spark: Spark 2.0 tools include Wikis, Blogs, Podcast Publishing, Forums, Media Markups for web- based video annotation, Google Map collections, and Web Meetings (<http://spark.uit.tufts.edu/>). The Spark wiki site will be made public to allow access to non-Tufts users and the link to this website will be posted prominently on the websites of all project senior personnel. The web-accessible data-sharing platform will include an easily accessible index listing where the data are stored and how they can be accessed. This index will be stored electronically with the actual files. We will additionally share data through publication in open-access journals where suitable and presentations at conferences, workshops, seminars and public lectures.

The farmer and consumer data products consist of human subjects research and (are) subject to IRB and privacy requirements. The raw data cannot be made publicly available. However the results will be published in peer-reviewed journals and presented.

|  |  |  |
| --- | --- | --- |
| **RESEARCH PRODUCTS** | **FORMAT OF DATA STORAGE** | **FORMAT OF DATA SHARING** |
| **A. Climate Variability** |  |  |
| Trends in precipitation and circulation over the study site | Atmospheric fields will be stored in netcdf format, one of the standard formats typically used in the ocean and atmospheric sciences  | Netcdf format allows 4-dimensional gridded data to be stored with all of the metadata; Web-accessible catalogue |
| Statistical output from a classification and regression tree (CART) model examining the variability of climate and tea prices | Excel spreadsheets; SPSS software output  | Excel spreadsheets on web-accessible catalogue; Write-up and graphs in scientific articles |
| Maps of climate trends | Digital .jpeg image files | Digital .jpeg image files on Web-accessible catalogue |
| **B. Secondary Metabolites of Leaves from *In Situ* Stratified Sampling and Manipulative Experiments** |  |  |
| Chromatogram and peak quantification output from secondary metabolite analysis  | EMPOWER, LCLab and ChemStation software data analysis output; Excel spreadsheets  | Chemical library maintained on web-accessible catalogue; Write-up and graphs in scientific articles |
| Laboratory Notes | Hardcopy notebooks; Digital copy of scanned notebook pages | Digital .jpeg filed of scanned notebook pages on web-accessible catalogue |
| **C. Farmer and Consumer Surveys** |  |  |
| Participant observation field notes  | Digital copy of scanned notebook pages; Typed notes in word processing files | Collection notebooks and all datasets will be provided to Smithsonian Institution as part of their efforts to gather legacy data related to ethnobiology and ethnoecology. |
| Questionnaire responses | Digital voice recordings; Typed notes in word processing files; Excel spreadsheets; Anthropac software data analysis output; SPSS software output | Ethnographic narrative on web-accessible catalogue; write-up and graphs in scientific articles  |
| Social network data  | Typed notes in word processing files; Excel spreadsheets; UCInet software data analysis output; .jpeg images | Digital .jpeg images on web-accessible catalogue; Write-up and graphs in scientific articles |
| Voucher specimens | Voucher specimens with typed descriptions using botanical standards distributed to University of Florida (FLAS), Smithsonian (US) and Missouri Botanical Garden (MO); Scanned voucher specimens  | Digital .jpeg images on web-accessible catalogue international biodiversity informatics websites |
| Prior informed consent forms | Forms will be scanned  | Not publically available as subject to privacy requirements |

**2. Re-Use, Re-Distribution and Production of Derivatives:** The PI will retain principal legal rights to the intellectual property developed under this grant as is compliant with the [Tufts University’s Policy on Rights and Responsibilities with Respect to Intellectual Property](https://exchange.tufts.edu/owa/redir.aspx?C=2b103aea43f44e1e9c706488ee060ad0&URL=http%3a%2f%2fviceprovost.tufts.edu%2f%3fpid%3d6%26c%3d6). Others, such as academic audiences, public policy practitioners, and local and state government are welcome to re-use or re-distribute our data with the only caveat that they identify us as the source. Data will also be made available on distribution media through requests made to PI. Minor handling fees will be charged as needed for such requests to support the costs of media and postage. PI will request appropriate acknowledgment in all publications or other use of data.

**3. Archiving Data:** All the collected data will be archived on the Tufts University Information Technology (UIT) research storage that is provisioned on Tufts’ NetApp enterprise networked storage. Data security and confidentiality are protected by using Microsoft Active Directory authentication, and the storage is backed up to LTO-4 tape on a daily and weekly basis and stored offsite at Iron Mountain facilities. As storage needs for the research study increase over time, additional storage can be easily provisioned on the same appliance. UIT issues storage space in 50GB increments, up to several TB. Only the latest data (1 year) will be made available online through the project website that will contain an easily accessible index to navigate the data and listing how they can be accessed while older data archived in the UIT research storage will be made available online for downloading in manageable sizes (~10GB) upon request. Updating the data on the web site is expected to be done every two weeks.