Utah State University

DigitalCommons@USU

Funded Research Records   Data Services

5-28-2021

REU Site: STEM for Plant Health

David W. Britt
Utah State University, david.britt@usu.edu

Follow this and additional works at: https://digitalcommons.usu.edu/funded_research_data

Part of the Biological Engineering Commons

Recommended Citation

Utah State University
MERRILL-CAZIER LIBRARY

4. Data Management Plan

a. Expected Data Type. The proposed REU primarily explores plant-microbe interactions through an undergraduate outreach program. All data are primary data. Some data such as recording plant height, weight, water use, are non-digital whereas analytical results are digital. Essential metadata will be recorded as headers in instrument software with associated data or in laboratory notebooks along with hand-recorded data. Hand-recorded data will be immediately transferred for storage on desktop computers. Laboratory notebooks are reviewed with the REU participants and peer mentors, and signed weekly by the Faculty Mentors.

For the program assessment and evaluation, metadata will include the self-efficacy instrument entered in a statistical analysis package and secondary products will include interviews in the form of audio files (MP4 format) and written transcripts (DOCX format).

b. Data Format. Raw and processed data and graphic representation of the data will be in Excel software, statistical analysis will be performed using R or JMP software, and imaging as jpeg files will be stored on the site for all to review. Each data file will be associated with a metadata file containing a full description of how the data were generated (experiment design, experimental variables, number of replicates, observations, QA/QC) so that all participants can access the data. All data will be backed up weekly on external drives and stored off-site using box.usu.edu, which provides unlimited storage space to faculty. All laboratory notebooks will be scanned on a weekly basis and archived. Final archiving of written material will be with the PIs. Findings will be compiled into Power Point presentations given to the group in designated weekly meetings of the PIs and all researchers and amended based on questions and comments. In addition, the mentors for undergraduate and graduate research will meet at other times individually to examine the notebooks with raw data, discussion of data handling and to ensure progress in research writing.

For the program assessment and evaluation, all de-identified files and participant consent forms will be kept in a cloud based Box folder only accessible to the PIs and Co-Is.

c. Data Storage and Preservation. All data will be stored in an encrypted cloud-based server (box.usu.edu) with access by all of the research group members. This research team has used this site for over two years and it is approved as a secure mechanism within Utah State University. Hard copies of data will be destroyed immediately after identifiers have been removed. A secondary copy of these data will be kept on a password-protected hard drive in a locked cabinet in a locked office in a location that is safeguarded against fire and water damage. Five years after the completion of the project, any items with identifiers (e.g., consent forms) will be destroyed. De-identified data will remain in the hard-drive for five years after the project has concluded to enable future analyses.

d. Data Sharing and Public Assess. Data will be placed on a shared network server using existing hardware at the Utah Water Research Laboratory for access to other consumers. Here we will publish our data files in a commonly supported CSV format and will describe them using compliant metadata files.

According to USU's Institutional Review Board (IRB), other researchers can view the data after the project is over and after identifying information has been removed. Although we will not provide open access to the participant data, the PI can make it available upon written request (and permission) to any researchers who might want to view it one year after the project phases pertinent to the data has been completed and/or the information has been analyzed, de-identified and submitted for publication. All identifying information will be removed before the data is made

available to any researcher and all data to be used in the workshops will be done in composite form with non-identifiable information. The results of data analyses will be shared through publications and conference presentations, but all participant identifiers will be removed from all data excerpts prior to publication. In the unlikely case that the journal editors request access to raw data or metadata, the PI will make the following items available to them: (a) tabular or graphical comparisons of participant entries; and (b) raw de-identified data, such as REU session transcripts.

e. Roles and Responsibilities. PI Britt will ensure that the DMP in fully implemented. The PIs have successfully managed USDA and NSF funded projects. PI Britt will serve as the QA officer. The PIs and Co-Is on this project are fully vested in supporting the DMP and would continue to implement the plan regardless of personnel changes. No addition resources are needed to support this DMP.

Quality assurance/quality control: We have an established QA/QC plan which contains the basic components of accuracy, precision, completeness, representativeness, and comparability as described in Test Methods for Evaluating Solid Waste, SW-846 (USEPA 1986 and updates); Guidance for the Data Quality Objectives Process (USEPA, QA/G4, 2000); EPA Guidance for Quality Assurance Project Plans (USEPA, QA/G-5, 1998) and Practical Methods for Data Analysis (USEPA QA/G-9, 2000). The raw data, and precision and accuracy checks on the raw data for the biological studies and chemical analyses, will be organized in a form to permit defense of the data and the conclusions drawn. Appropriate procedures will be utilized to ensure that all samples collected for analysis are statistically representative of the studied system. Use of standardized methods of analysis will ensure the comparability of results. Standardized data format for collection, calculation and reporting of data will facilitate the generation of comparable data.

Data analysis will employ statistical programs, such as R and JMP. Statistics to be used will include regression analysis, descriptive statistics such as the mean, range, and confidence intervals for the mean. Comparative and inferential data analysis will be done using analysis of variance procedures and/or multiple linear regression analyses. Exploratory procedures, such as principal component analyses, may also be used. The data reduction schemes for analytical measurements, including all equations used to calculate concentration or values of measured parameters and reporting units are contained in standard methods or referred publications. All experiments will be replicated at least three times. Appropriate control samples will be employed.

6. Monitoring and Reporting
Monitoring of our programmatic objectives will take place through our external evaluator, Co-I. Villanueva, who will assist us in identifying that the goals of the project are being met as well as provide yearly reports regarding student outcomes. Furthermore, yearly reflective focus groups will occur between the PIs and the student facilitators to ensure that any challenges from the prior REU can be addressed in a timely manner. All findings from the program evaluator and from our meetings, including our publications and findings to date will be assembled into annual reports to NSF. De-identified raw data will be stored for 5 years after the award ends.