FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

INTRODUCTION TO LSIL DATA MANAGEMENT

SETTING YOUR PROJECT UP FOR

Feed the Future Innovation Lab for Livestock Systems

Lacey Harris-Coble

USAID FROM THE AMERICAN PEOPLE

BILL & MELINDA GATES foundation

ILRI INTERNATIONAL LIVESTOCK RESEARCH INSTITUTE

CGIAR

UF IFAS UNIVERSITY of FLORIDA

PURPOSE & TARGET AUDIENCE

Objective:

Understand and manage data reporting requirements to fulfill USAID's Open Data Policy

Target audience:

- Principal Investigator

Also appropriate for:
Co-Principal Investigators
Researchers or Post-docs

USAID'S OPEN DATA POLICY

- USAID's open data policy (ADS 579) addresses both the legal and technical aspects of open data.
- With the issuance of ADS 579 (in 2010) and its associated changes to award terms, new and existing USAID awards now require USAID implementing partners to submit datasets generated with USAID funding to the Development Data Library (DDL) in **machine-readable, non-proprietary formats**.
- Datasets are published at www.usaid.gov/data, unless legally exempted, after undergoing an internal clearance process.
- Published datasets also appear automatically at https://data.gov/

Check out the types of data available on the Data Development Library for your country

USAID'S OPEN DATA POLICY

- **Datasets submitted to the DDL must also be accompanied by supporting documentation**, defining the fields within the Dataset and any categories or labels within the Dataset that may require explanation to an individual not familiar with the data.
- This may be accomplished through the submission of a codebook or data dictionary. When available, scopes and methodologies, such as survey protocols and instruments used to collect and analyze the data must also be submitted to the DDL, along with annotations to inform the general public of any known data quality issues.
- Find out more info on Frequently Asked Questions About Open Data

HOW DOES LSIL'S DATA MANAGEMENT WORK?

LSIL's data management is comprised of two parts:

1. Datasets Module in Piestar
2. Dataset Submission to the DDL

HOW DOES LSIL'S DATA MANAGEMENT WORK?

LSIL's data management is comprised of two parts:

1. Datasets Module in Piestar
   - **When?** Beginning of project, updated on annual basis
2. Dataset Submission to the DDL

HOW DOES LSIL'S DATA MANAGEMENT WORK?

LSIL's data management is comprised of two parts:

1. Datasets Module in Piestar
   - **When?** Beginning of project, updated on annual basis
   - **PI Role:** Enter and update datasets planned in project activities
2. Dataset Submission to the DDL

HOW DOES LSIL'S DATA MANAGEMENT WORK?

LSIL's data management is comprised of two parts:

1. Datasets Module in Piestar
   - **When?** Beginning of project, updated on annual basis
   - **PI Role:** Enter and update datasets planned in project activities
   - **LSIL Role:** Collate dataset information from all projects and submit to USAID
2. Dataset Submission to the DDL

HOW DOES LSIL'S DATA MANAGEMENT WORK?

LSIL's data management is comprised of two parts:

1 Datasets Module in Piestar

- **When?** Beginning of project, updated on annual basis
- **PI Role:** Enter and update datasets planned in project activities
- **LSIL Role:** Collate dataset information from all projects and submit to USAID

2 Dataset Submission to the DDL

- **When?** At the end of project, after final report submission

HOW DOES LSIL'S DATA MANAGEMENT WORK?

LSIL's data management is comprised of two parts:

1 Datasets Module in Piestar

- **When?** Beginning of project, updated on annual basis
- **PI Role:** Enter and update datasets planned in project activities
- **LSIL** Role: Collate dataset information from all projects and submit to USAID

2 Dataset Submission to the DDL

- **When?** At the end of project, after final report submission
- **PI Role:** send all data files indicated in Open Data plan along with supporting documents to LSIL

HOW DOES LSIL'S DATA MANAGEMENT WORK?

LSIL's data management is comprised of two parts:

1 Datasets Module in Piestar

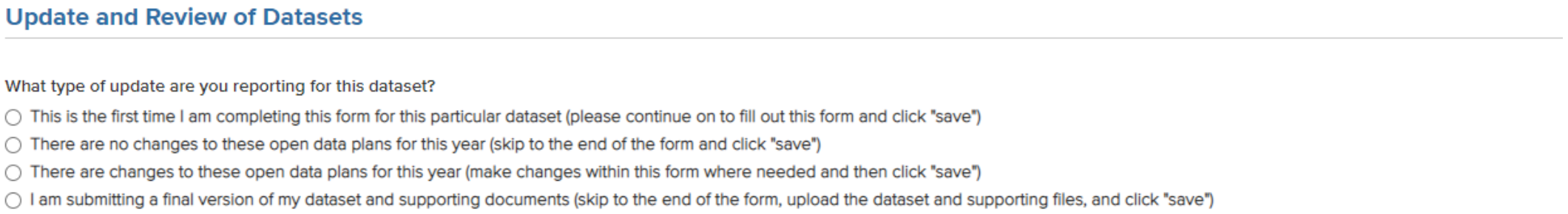
- **When?** Beginning of project, updated on annual basis
- **PI Role:** Enter and update datasets planned in project activities
- **LSIL Role:** Collate dataset information from all projects and submit to USAID

2 Dataset Submission to the DDL

- **When?** At the end of project, after final report submission
- **PI Role:** send all data files indicated in Open Data plan along with supporting documents to LSIL
- **LSIL Role:** review and ensure files meet DDL criteria, submit to DDL on behalf of project

WHAT DO YOU NEED TO KNOW TODAY?

- Complete the Piestar Datasets Management module
  - ❖ If you plan to collect qualitative datasets, please include them in the Datasets module, but note that they do not have to be submitted to the DDL at the end of the project.
  - ❖ Update information annually, as needed



FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

COMPLETING THE DATASETS PIESTAR MODULE

Dataset Details

Please provide a name for this dataset.

Provide details to describe the dataset. Please answer the following questions: What information will be collected? Where will the information be collected? When will the data be collected?

Please provide 3 to 5 key words to describe the dataset.

FEED THE FUTURE
The U.S. Government's Global Hunger & Food Security Initiative

COMPLETING THE DATASETS PIESTAR MODULE

Data File Type and Supporting Documents

Which of the type of data listed best describes this dataset?

- ◯ Genomic sequence data
- ◯ Molecular marker data
- ◯ Crop phenotypic data from on-station field trials
- ◯ Crop phenotypic data from farmer-managed field trials
- ◯ Set of crop-specific phenotyping protocols
- ◯ Farmer survey data
- ◯ Anthropometric/health data from subjects
- ◯ Seed value chain/local market data
- ◯ Pest/pathogen phenotypic data
- ◯ None of the above

Describe the estimated total number of observations the dataset will include and the type of file that will be used (e.g. Excel file with 180 observations). The final data must be in a non-proprietary machine-readable format (e.g. csv file).

Indicate the types of supporting documents that will be submitted along with the data file. Please note that codebooks are required for all data files. Please see the Piestar Dashboard resources for an example of the type of codebook that is required. IRB approval, informed consent scripts and survey instruments are relevant for data collected from human subjects.

- ☐ Codebook (required)
- ☐ IRB approval
- ☐ Informed consent script
- ☐ Survey instrument or questionnaire
- ☐ Other

COMPLETING THE DATASETS PIESTAR MODULE

Privacy Concerns

Describe any personally identifiable information (e.g. age, religion, race/ethnicity) or contact information (e.g. name, phone number, address) that will be part of this dataset.

Describe any privacy concerns or restrictions on making the data public.

To ensure that data quality & data privacy concerns are addressed before making it publicly available, describe any data-processing steps (cleaning, making responses anonymous, changing data to represent a higher geospatial scale, etc.) that the project will perform on the raw dataset.

COMPLETING THE DATASETS PIESTAR MODULE

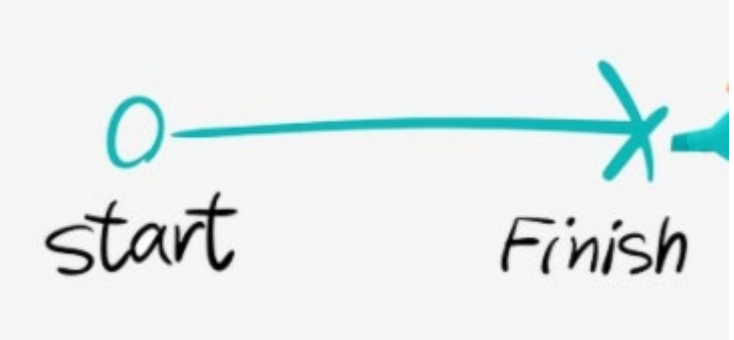
Embargo Period and Submission

Who will be responsible for ensuring data to be submitted? In most cases this is the PI. If the proposed person is not the PI, please provide an explanation.

Will you request an embargo period for this dataset? That is, will you request a period of time in which the data is not published online after you submit your final data files? If so, please indicate how long you are requesting an embargo.

LOOKING AHEAD – START WITH THE END IN MIND

1. From the start, organize and save/safeguard data files and supporting documents (e.g., codebooks, informed consent scripts, questionnaires etc.) in a way that will make it easy to do the dataset submission at the end of your project!



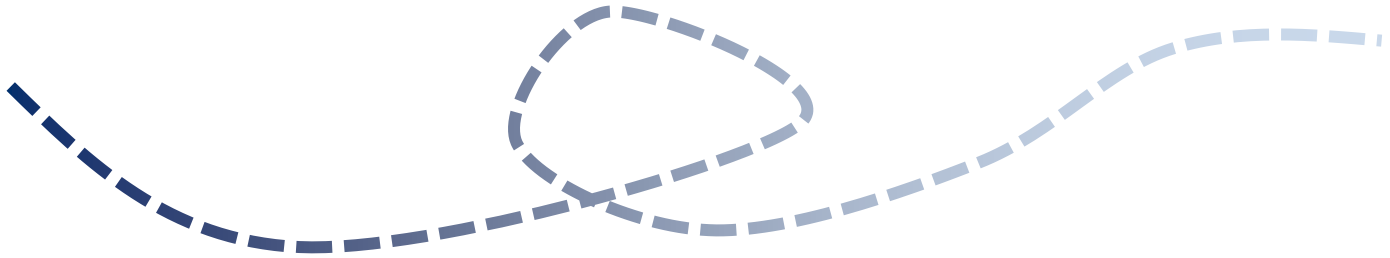
2. Take the forthcoming self-paced UF Canvas course later this year

   - It provides guidance on how to prepare datasets for submission to the DDL and include tips, e.g., on how to ensure data collected via tablets is structured properly for data submission.

THANKS FOR YOUR TIME AND ATTENTION!

Remember:

When in doubt – reach out!



LSIL

Disclaimer

This work was funded in whole or part by the United States Agency for International Development (USAID) Bureau for Food Security under Agreement # AID-OAA-L-15-00003 as part of Feed the Future Innovation Lab for Livestock Systems. Any opinions, findings, conclusions, or recommendations expressed here are those of the authors alone.

Feed the Future Innovation Lab for Livestock Systems
Department of Animal Sciences | University of Florida | P.O. Box 110910 | Gainesville, FL 32611
livestock-lab@ufl.edu | http://livestocklab.ifas.ufl.edu/

FEED THE FUTURE

The U.S. Government's Global Hunger & Food Security Initiative

www.feedthefuture.gov

USAID FROM THE AMERICAN PEOPLE

BILL & MELINDA GATES foundation

ILRI INTERNATIONAL LIVESTOCK RESEARCH INSTITUTE

CGIAR

UF IFAS UNIVERSITY of FLORIDA