

# Texas Data Repository Tutorial

## ILSSI Dataset Submission

## About the Texas Data Repository

The Texas Data Repository (TDR) is an online platform for Texas A&M University researchers to publish and archive datasets and data products. It is accessed as a website.

The TDR is built on Dataverse, an open-source software developed by Harvard University. It is hosted by the Texas Digital Library, a consortium of Texas academic libraries focused on long-term access and preservation of digital content. The TDR is committed to preserving and providing ongoing access to research data for at least 10 years after submission.

More information about the TDR is available here: <http://data.tdl.org/about/>

### Datasets

In the Texas Data Repository, data files are uploaded to “datasets.” Each dataset receives its own webpage and a Digital Object Identifier (DOI).

A dataset may hold multiple files, these files can contain data or documentation about the data. Each file within a dataset may be **2GB** or less and should be saved in a **non-proprietary** file format, for example a comma-separated value file (CSV) instead of a Microsoft Excel Spreadsheet file (XLSX). The files uploaded to a dataset can also be ZIP or TAR files that contain multiple files within them.

Data file(s) uploaded to a dataset are stored along with additional information about the data, such as the title, creator, and date created. This “metadata” is structured, and describes the dataset to users and the Texas Data Repository system in a consistent way. It is visible on the dataset webpage and used by the Texas Data Repository software for search and filter features.

### Dataverses

“Dataverse” is both the name of the software running the Texas Data Repository and the name of a container that holds datasets. The dataverse container is similar to a directory or folder in a computer file system.

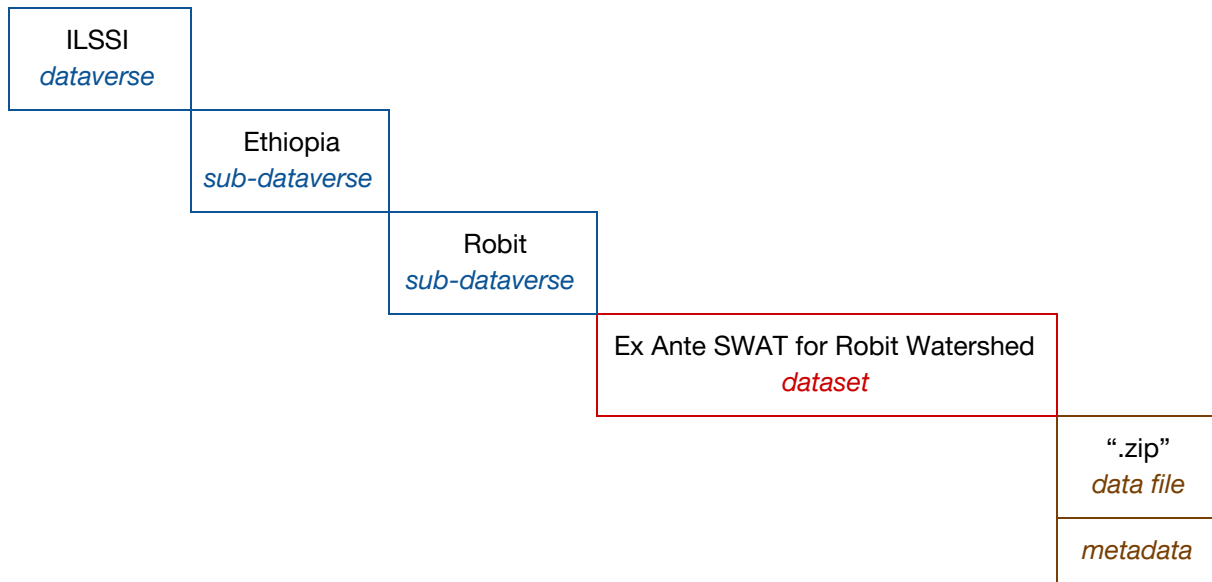
A dataverse may hold one or more datasets, it can be used to organize datasets into collections based on a project or other criteria. The dataverse container can also hold sub-dataverses, that is, sub-collections of datasets.

## ILSSI Dataverse

The ILSSI dataverse is organized by country and region. It contains sub-dataverses named after countries, these country dataverses contain sub-dataverses named after regions, they may also hold datasets that apply to the country level.

Each regional sub-dataverse includes datasets relevant to that region. Each dataset may be based on a different model and analysis conducted in a particular region.

### Example of Structure



### Permissions

Researchers are given permissions to create datasets within a particular dataverse by the ILSSI Program Coordinator, or alternatively, the [Data Management Librarian](#).

## Adding Datasets

Access the ILSSI Dataverse

Go to: <https://dataverse.tdl.org/dataverse/ilssi>

## Log In

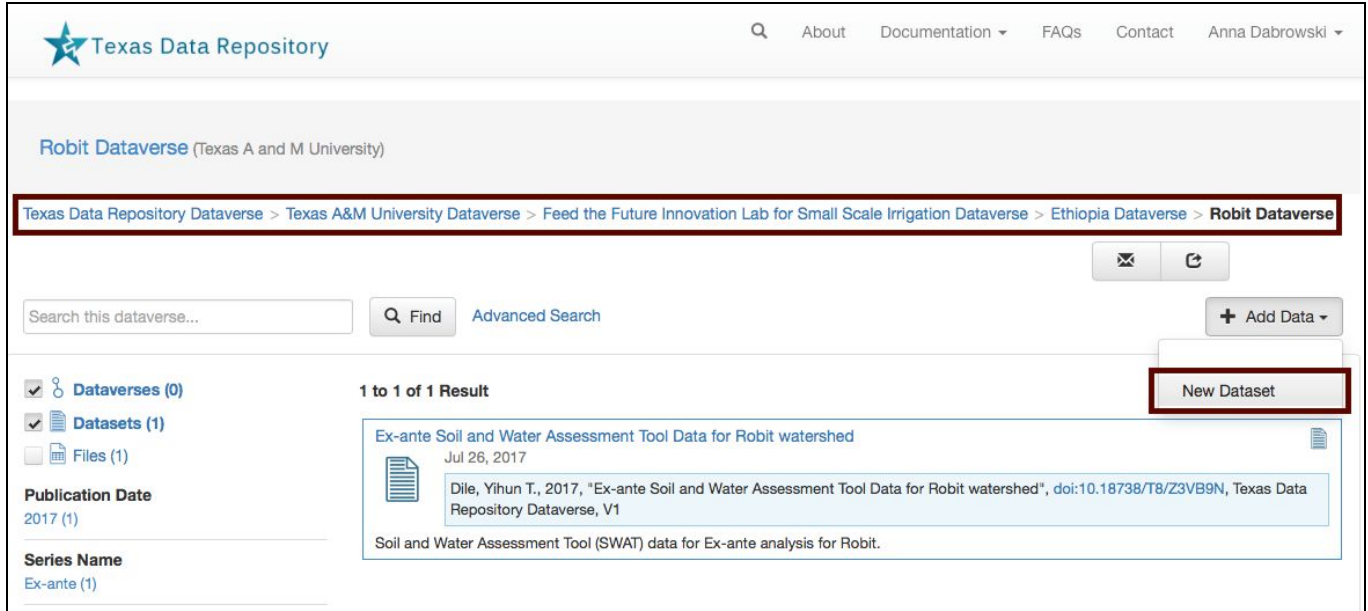
Click “Log In” at the upper right.

For researchers at Texas A&M University, choose “Texas A and M University.” Use your TAMU NetID and password to log in.

### Create a Dataset

Go to the dataverse for the region where you would like to submit a data.

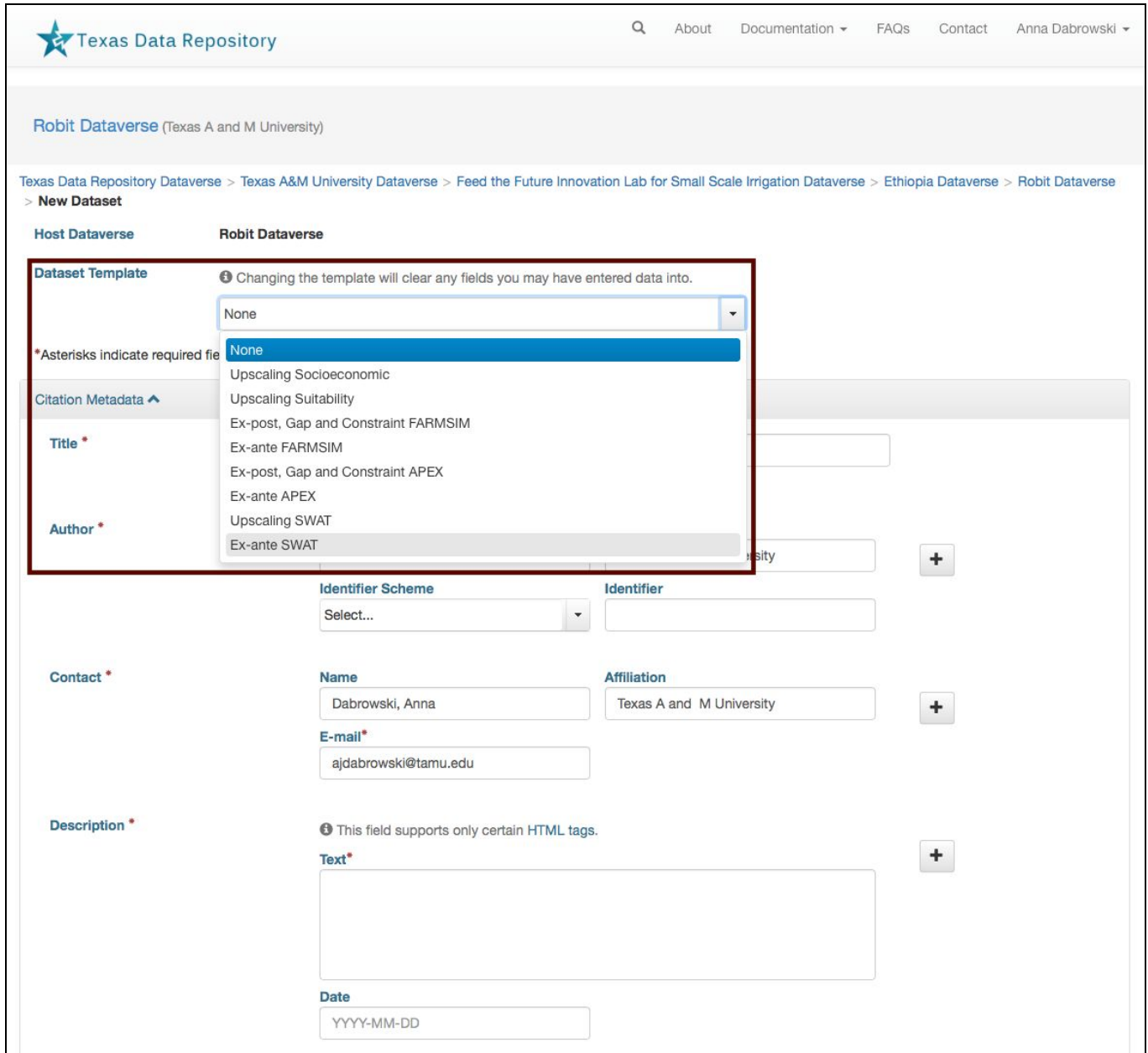
Choose “New Dataset.”



The screenshot shows the Texas Data Repository website interface. At the top, there is a navigation bar with the Texas Data Repository logo and links for About, Documentation, FAQs, Contact, and a user profile for Anna Dabrowski. Below the navigation bar, the breadcrumb trail is highlighted in red: **Texas Data Repository Dataverse > Texas A&M University Dataverse > Feed the Future Innovation Lab for Small Scale Irrigation Dataverse > Ethiopia Dataverse > Robit Dataverse**. The main content area features a search bar with the text "Search this dataverse...", a "Find" button, and an "Advanced Search" link. To the right of the search bar are buttons for "Add Data" and "New Dataset", with the latter being highlighted in red. On the left side, there are filters for "Dataverses (0)", "Datasets (1)", and "Files (1)", along with sections for "Publication Date" (2017 (1)) and "Series Name" (Ex-ante (1)). The main results area shows "1 to 1 of 1 Result" for the dataset "Ex-ante Soil and Water Assessment Tool Data for Robit watershed", dated Jul 26, 2017. The dataset description includes the author "Dile, Yihun T., 2017, 'Ex-ante Soil and Water Assessment Tool Data for Robit watershed', doi:10.18738/T8/Z3VB9N, Texas Data Repository Dataverse, V1" and the description "Soil and Water Assessment Tool (SWAT) data for Ex-ante analysis for Robit."

## Add Metadata

Choose a “Dataset Template” named after the model and analysis used for your data.



The screenshot shows the 'New Dataset' form in the Texas Data Repository. The 'Dataset Template' dropdown menu is open, showing options: None, Upscaling Socioeconomic, Upscaling Suitability, Ex-post, Gap and Constraint FARMSIM, Ex-ante FARMSIM, Ex-post, Gap and Constraint APEX, Ex-ante APEX, Upscaling SWAT, and Ex-ante SWAT. The form includes fields for Title, Author, Identifier Scheme, Identifier, Contact Name, Affiliation, E-mail, Description, and Date.

**Dataset Template** ⓘ Changing the template will clear any fields you may have entered data into.

None

\*Asterisks indicate required fields

Citation Metadata ▲

**Title \***

**Author \***

**Identifier Scheme** Select... **Identifier**

**Contact \***

**Name** Dabrowski, Anna **Affiliation** Texas A and M University

**E-mail \*** ajdabrowski@tamu.edu

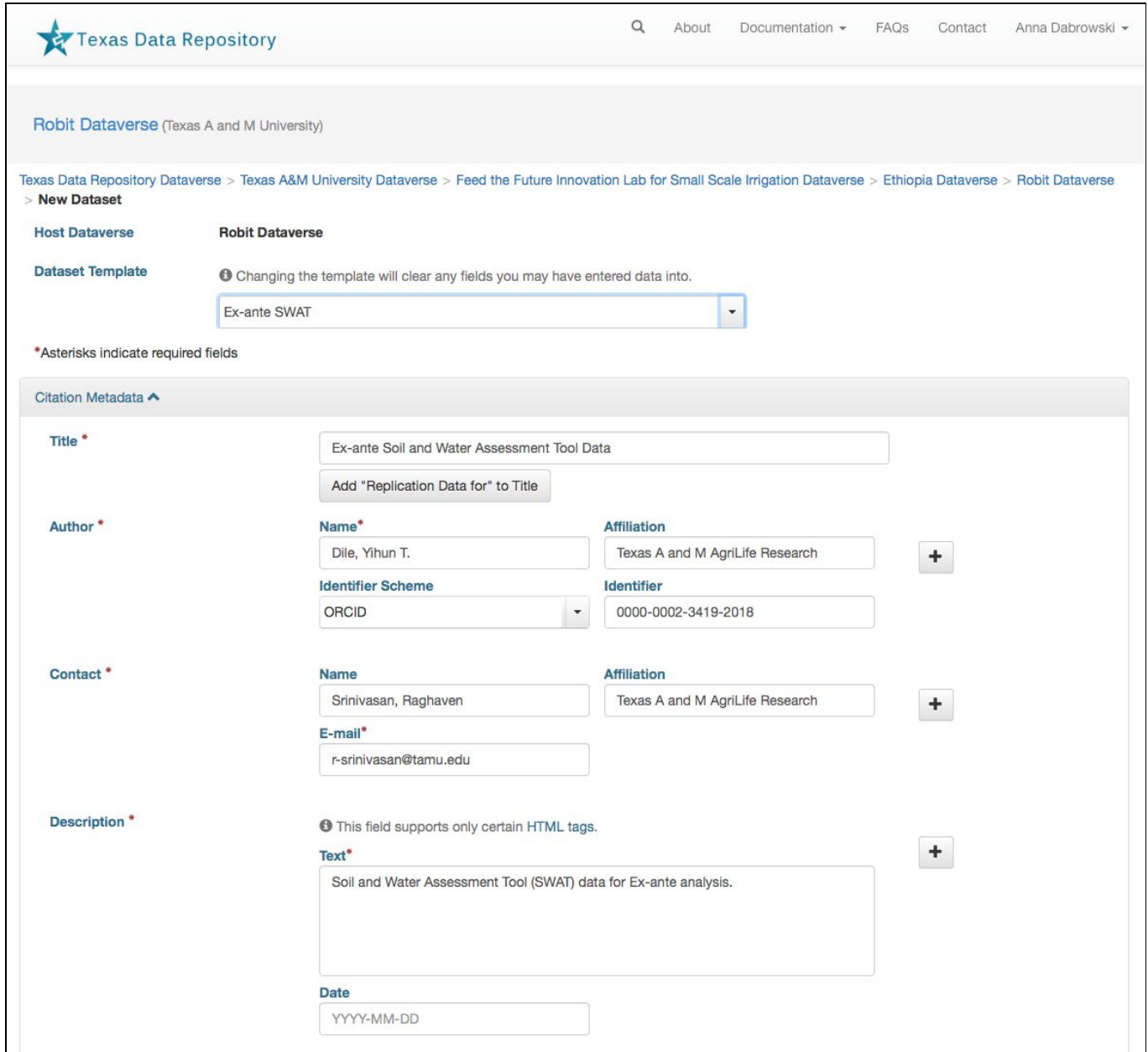
**Description \*** ⓘ This field supports only certain HTML tags.

**Text \***

**Date** YYYY-MM-DD

**N.B.** The templates will automatically fill descriptive information for you.

If a template is incorrect, please contact the ILSSI Program Coordinator or [Data Management Librarian](#) to change the template.



**Texas Data Repository** | About | Documentation | FAQs | Contact | Anna Dabrowski

**Robit Dataverse** (Texas A and M University)

Texas Data Repository Dataverse > Texas A&M University Dataverse > Feed the Future Innovation Lab for Small Scale Irrigation Dataverse > Ethiopia Dataverse > Robit Dataverse > **New Dataset**

**Host Dataverse:** Robit Dataverse

**Dataset Template:** Ex-ante SWAT

\*Asterisks indicate required fields

**Citation Metadata**

**Title \***  
Ex-ante Soil and Water Assessment Tool Data  
Add "Replication Data for" to Title

**Author \***

Name *	Affiliation
Dile, Yihun T.	Texas A and M AgriLife Research

**Identifier Scheme:** ORCID | **Identifier:** 0000-0002-3419-2018

**Contact \***

Name	Affiliation
Srinivasan, Raghaven	Texas A and M AgriLife Research

**E-mail \*:** r-srinivasan@tamu.edu

**Description \***  
This field supports only certain HTML tags.  
**Text \*:** Soil and Water Assessment Tool (SWAT) data for Ex-ante analysis.

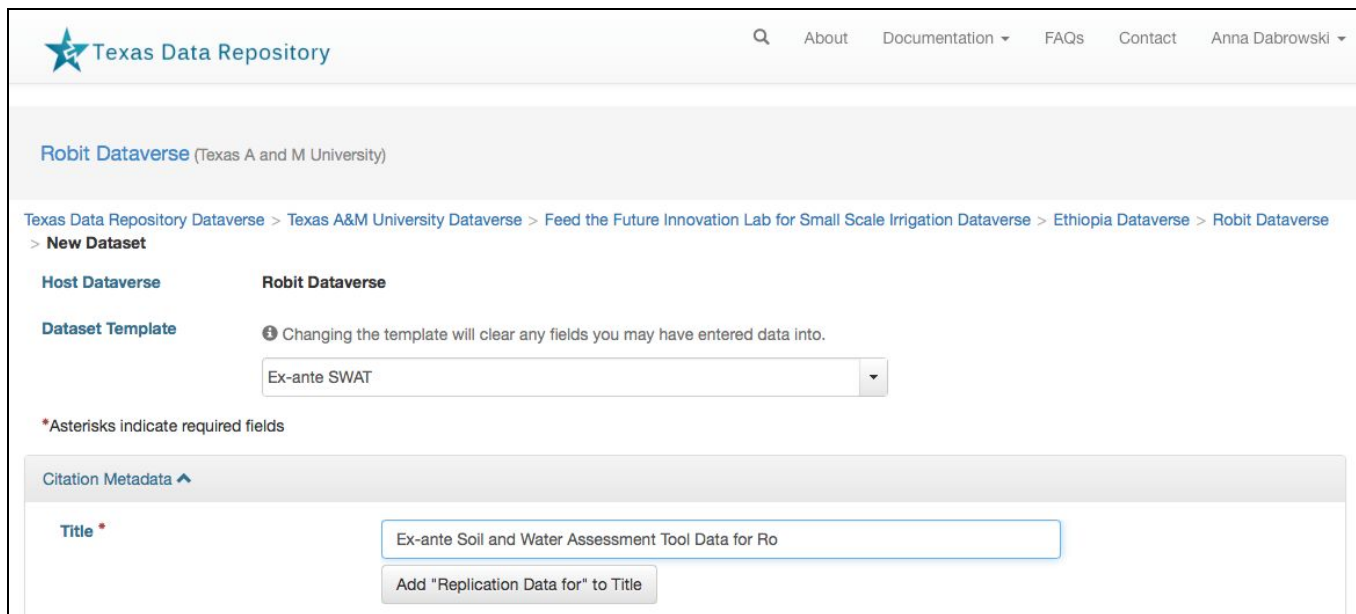
**Date:** YYYY-MM-DD

You will need to add some information:

1. Change the title to include the location
2. Choose the correct country
3. Add geographic state/region

**N.B.** Be consistent in capitalization, spelling, terminology, and the format of this information.

1. Change the title to include the location:



Texas Data Repository

Robit Dataverse (Texas A and M University)

Texas Data Repository Dataverse > Texas A&M University Dataverse > Feed the Future Innovation Lab for Small Scale Irrigation Dataverse > Ethiopia Dataverse > Robit Dataverse > **New Dataset**

**Host Dataverse** Robit Dataverse

**Dataset Template** Changing the template will clear any fields you may have entered data into.

Ex-ante SWAT

\*Asterisks indicate required fields

Citation Metadata

**Title \*** Ex-ante Soil and Water Assessment Tool Data for Ro

Add \*Replication Data for\* to Title



2. Choose the correct country:

**Geographic Coverage**

Country / Nation*	Ghana	State / Province*		+	-
City		Other			
Country / Nation*	Ethiopia	State / Province*		+	-
City		Other			

**Geographic Bounding Box**

West Longitude*		East Longitude*		+
North Latitude*		South Latitude*		

3. Add geographic state/region

**Related Datasets**

International Food Policy Research Institute (IFPRI) <a href="https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DH103J">https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DH103J</a>	+
---	---

**Geographic Coverage**

Country / Nation*	Ethiopia	State / Province*	Amhara	+
City		Other		

**Geographic Bounding Box**

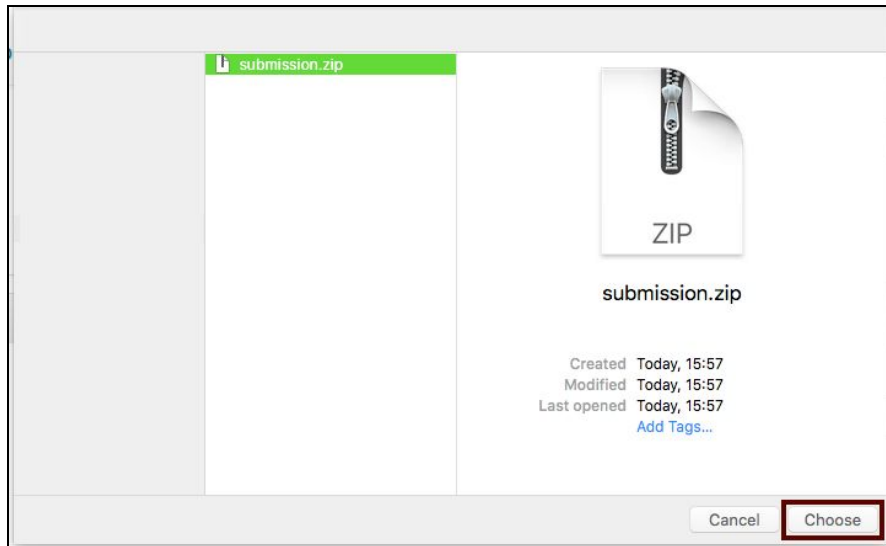
West Longitude*		East Longitude*		+
North Latitude*		South Latitude*		

### Upload Data Files

Make sure that each ZIP file is less than 2GB, split larger files into multiple ZIP files. Click “Select Files to Add”

**N.B.** If your ZIP file contains hundreds of files, the system will be unable to extract them. Dataverse will provide a warning message.

You may continue, this means that the repository will store your ZIP file, without un-zipping the contents.



Once your files have uploaded, click “Save Dataset” to submit the dataset for review.

**Dataset Version: DRAFT**


[Files](#)

File upload limit is 2,147,483,648 bytes per file. For more information about supported file formats, please refer to the [User Guide](#).

**+ Select Files to Add**

---

**1 File**

<input type="checkbox"/>	<input type="checkbox"/>	<b>File Name</b> <input type="text" value="submission.zip"/>	
		ZIP Archive MD5: 1a9b361ff1e78467f99303c4f1c71919;	<input type="button" value="Edit Tags"/>
	<b>Description</b>	<input type="text" value="Add file description..."/>	

Metadata Tip: After adding the dataset, click the Edit Dataset button to add more metadata.

Your dataset will be “in review” and “unpublished” until an administrator reviews and publishes it. Once it is published, the dataset will be publically visible and the files will be openly accessible for download.

The screenshot shows the Texas Data Repository interface. At the top, there is a navigation bar with the repository logo and links for About, Documentation, FAQs, Contact, and a user profile for Anna Dabrowski. Below this is the breadcrumb trail: Texas Data Repository Dataverse > Texas A&M University Dataverse > Feed the Future Innovation Lab for Small Scale Irrigation Dataverse > Ethiopia Dataverse > Robit Dataverse. The dataset title is "Ex-ante Soil and Water Assessment Tool Data for Robit watershed" with 0 Downloads. A description box contains the citation: "Dile, Yihun T., 2017, 'Ex-ante Soil and Water Assessment Tool Data for Robit watershed', doi:10.18738/T8/Z3VB9N, Texas Data Repository Dataverse, V1". Below the description is a table with fields for Description, Subject, and Keyword. The Description field contains "Soil and Water Assessment Tool (SWAT) data for Ex-ante analysis for Robit." The Subject field contains "Agricultural Sciences" and the Keyword field contains "IDSS, Hydrology, Environmental, Agriculture, SWAT, Irrigation". There is a "Dataset Version: 1.0" section with tabs for Files, Metadata, Terms, and Versions. A search bar is present with the text "Search this dataset...". Below the search bar, it shows "1 File" and a list of files. The file list contains one entry: "Robit-TxtInOut.zip", which is a ZIP Archive of 43.6 MB, dated Jul 26, 2017, with 0 Downloads. The MD5 hash is 4f24fe6576003bd61661e2200e0152a5. There are buttons for "Download", "Upload Files", and "Edit Files". At the bottom of the page, there are logos for TDL.ORG (Texas Digital Library) and The Dataverse Project v. 4.6.1.

### Office of Scholarly Communications contact information

Contact us for information, consultations, and training.

**Anna J Dabrowski**  
 Data Management Librarian  
[ajdabrowski@library.tamu.edu](mailto:ajdabrowski@library.tamu.edu)  
 979.845.8847

**Dr. Bruce Herbert**  
 Director, Office of Scholarly Communications  
[beherbert@library.tamu.edu](mailto:beherbert@library.tamu.edu)  
 979.845.1083