Salmon Data Discovery Tool Data Import Guide

Contents

Introduction

The Salmon Data Discovery Tool (SDDT) is a joint effort of NOAA's Northwest Fisheries Science Center and the Pacific States Marine Fisheries Commission. This web-based data repository fills a need to publicly house salmon relevant data currently held by various organizations, co-managers, and research scientists. The data discovery tool provides a single location to store data across various ecosystems (freshwater, estuary/nearshore, off-shore ocean) and data types (documents, data files, derived data products, spatial data, and GitHub links). SDDT accepts both salmon run specific information (population abundance estimates, productivity trends, age data) and data related to environmental drivers (e.g., marine heat waves, abundance and survival rates of predator and prey species, freshwater habitat quality and quantity). SDDT serves as an information clearinghouse for a broad user community, ranging from casual data exploration and browsing to comprehensive data downloading supporting regional harvest and conservation management analysis.

For more information please contact SDDT@psmfc.org

This guide serves as a resource for data users to familiarize themselves with the data uploading process and learn to manage data submissions independently.

Why share data with SDDT?

- **Enhanced Data Utility**: Make data findable and shareable
- **Collaboration Opportunities**: Sharing data fosters collaboration among organizations, leading to potential joint projects, funding opportunities, and shared resources.
- **Resource Management**: A comprehensive data repository aids in effective fisheries management, conservation planning, and policy-making.

Preparing Your Data for Import

Preparing data before publishing to SDDT is essential for ensuring quality, protecting sensitive information, and maximizing usability. This involves cleaning and organizing the data, removing or anonymizing any personally identifiable information, and creating clear documentation such as metadata that explains variables, methods, and limitations.

Before uploading, ensure your data is:

- *Clean and validated*: Remove any corrupted or incomplete entries
- *Remove sensitive information* before uploading
- *Properly formatted*: Follow consistent date formats, number formats, etc.
- *Within size limits*: Check platform restrictions of 2 GB
- *Properly named*: Use descriptive, meaningful file names
- *Group related files* in logical collections
- *Use version numbers* for updated datasets
- *Compress the files into a ZIP file*

Supported File Formats

Common supported formats typically include:

- CSV files (.csv) - Comma-separated values
- Excel files (.xlsx, .xls) - Microsoft Excel spreadsheets
- XML files (.xml) - Extensible Markup Language
- Text files (.txt, .doc. .pdf) - Plain text data or Adobe Acrobat files
- Spatial files (.gdb or .shp) – Geodatabase or Shapefile. These files will most likely need to be zipped before uploading
- Access file (.accdb, .mdb) – Microsoft Access database
- ZIP File (.zip) – Compressed zip file

Metadata

We highly encourage uploading metadata associated with tabular or spatial data e.g. ESRI format metadata files can be uploaded with your data. If you need a metadata template, we provide one here:
https://sddt.psmfc.org/SDDT_Metadata_Template_for_Tabular_Data_2025.xlsx

Citation

Data citations and acknowledgements are essential for maintaining scientific integrity and enabling reproducible research. Properly citing datasets gives credit to the researchers who collected, curated, and shared the data, encouraging continued data sharing within the scientific community. Citations also provide crucial information needed to locate, access, and verify the exact datasets used in research, making studies reproducible and transparent.

We encourage you to upload a document with full citation information for your file.

The uploaded citation file will be included in the compressed dataset downloaded file.

Creating a User Account

User account, manage user profile, dataset uploads.

1. Click on the ‘Login’ button in the upper left on the homepage
2. In the Login page, select 'Sign Up' to create a new account
3. Enter your information applicable to each field and select 'Sign Up' (See figure 1)
4. An activation email will be sent to the email provided during in the Sign Up window, with an activation link. Check your email and select the link to activate your new account. If you did not get the email, check within your spam folder.

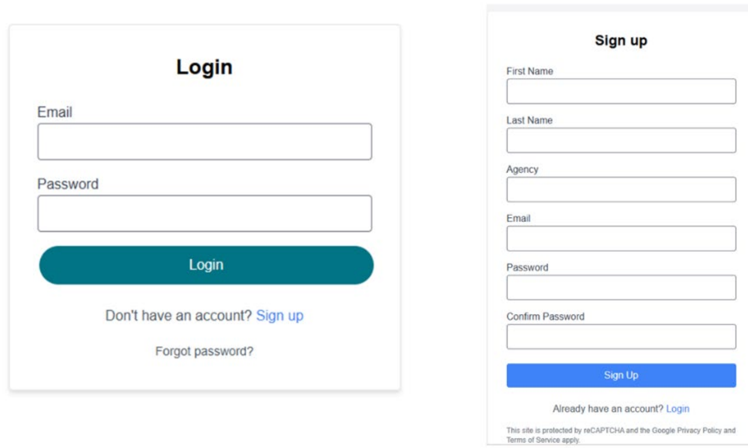


Figure 1: Login and New User account

User Profile

Once you have created your user profile, you can make any changes to your information or change your password. You can find your My Profile page under the drop-down menu on the Right.

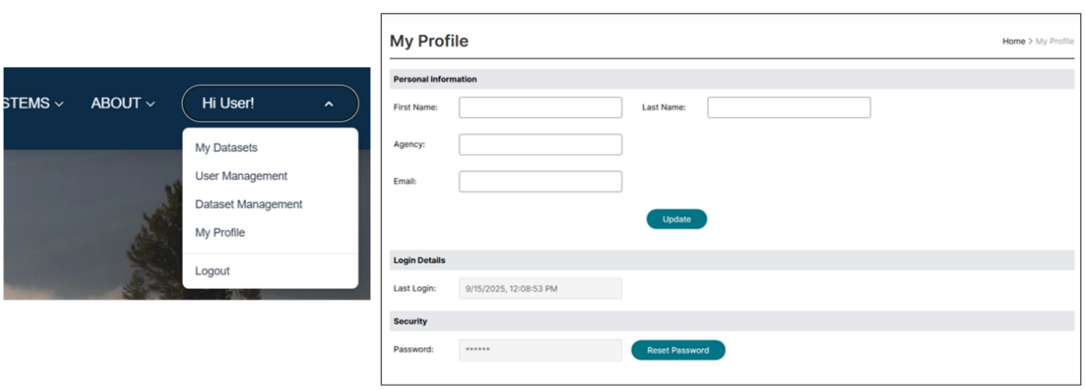


Figure 2: Example of the drop down menu and the My Profile page

Data Uploading Process

To add a new Dataset:

The data uploading page is broken down into five sections: Content Description, Citations, Contacts, Upload/Data Connect, Dataset Coverage Areas. This User Guide will only cover the required and recommended fields, but we highly suggest that you fill out as many of the data fields as possible to inform users about your file. To see the definitions of each data field, please visit the Controlled Vocabulary (weblink here).

1. Click on "Add New Dataset" in the top right corner of the screen (see figure 3).

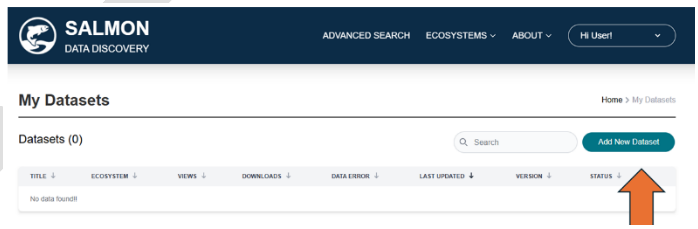


Figure 3: User account homepage defaults to the My Datasets Page

Content Description:

"Status" in the top left of the screen will be set to Draft by default. This will mean the dataset uploading process and information will be saved in your profile but not made public. It is advised you keep this option as Draft until you are ready to make your data public.

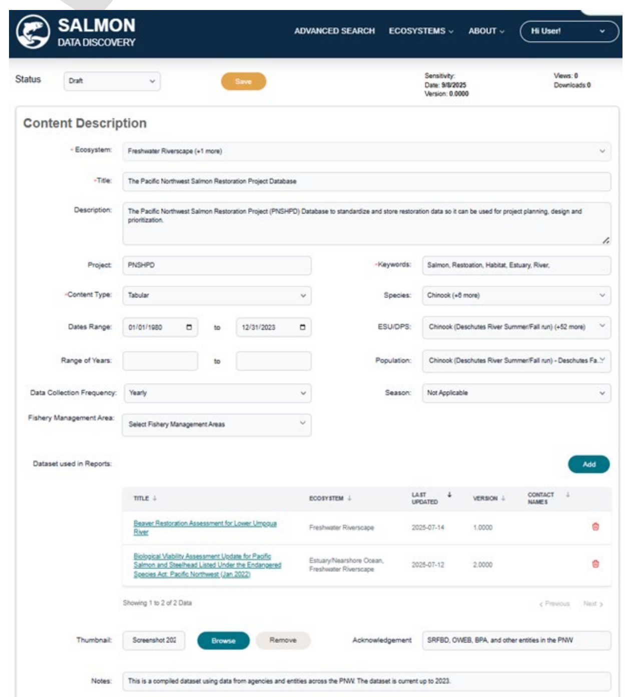


Figure 4: Example of the Content Description on the Data Uploading page.

1. *Ecosystem:* Select which ecosystem best represents your file. This is a multi-select option.
2. *Title:* The name of your file.
3. *Project:* If there is a research project, (another term) associated with the file you are uploading, type it here.
4. *Keywords:* Type in keywords that are associated with the file you are uploading. This field is one of the ways your file is searchable on SDDT. Please see the note above regarding Keywords
5. *Content Type:* This is a drop down menu to describe your file: Tabular, Spatial,

Document. For more details regarding file types, please visit the SDDT Controlled Vocabulary (provide link when ready)

6. *Fisheries Management Area:* These are the current, as of 2025, fishing zones listed by each state (on the West Coast) and the Pacific Fisheries Management Council) PFMC).
7. *Data Set Used in Reports:* This is where you can connect multiple datasets together so when they are searched, a user can see all the linked files in the search card. Click on the “Add” button on the right (see figure 5). A pop-up screen will appear with a list of Reports to link your file to. Using the check box to the left the name of the report, select which file you would like to link to. Click Save when you are done.

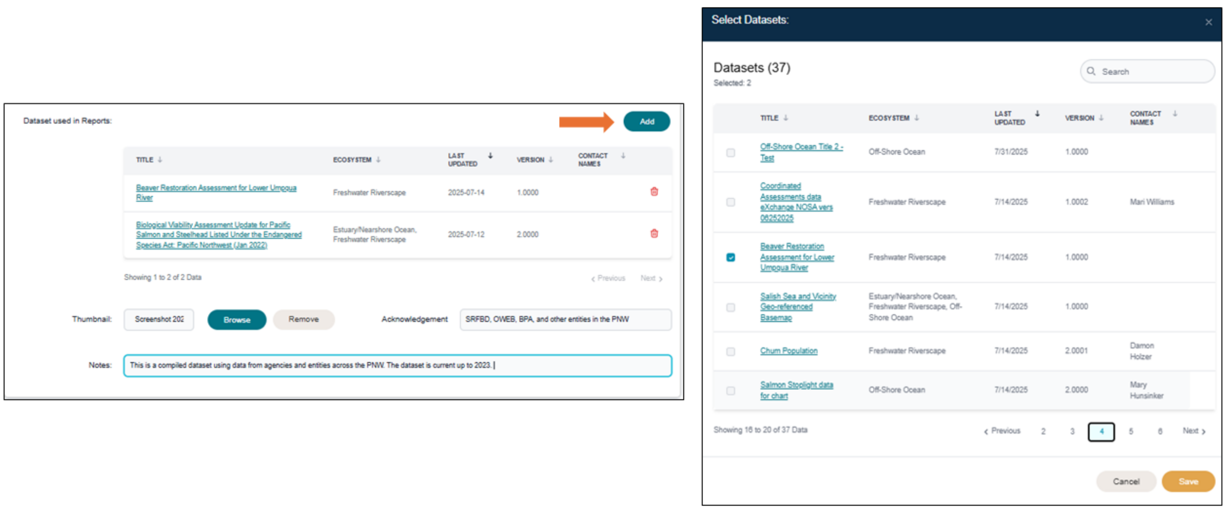


Figure 5: Image of the "Dataset Use in Reports” and Thumbnail sections of the Data Uploading Page and the pop-up screen with a list of reports.

Thumbnail: If you would like to upload an image to help users with a visual of the data, report, or a map of the spatial layer, you can upload it here but clicking the Browse button. Navigate to the image you would like to upload and select save (see figure 4). Recommend size for thumbnail: 150x100 pixels

Citations

Data citations and acknowledgements are essential for maintaining scientific integrity, giving proper credit to data creators, and enabling others to locate and verify datasets for reproducible research. While uploading citation information is encouraged but not required, completing this section helps users quickly identify who contributed to the data or documents being shared, promoting transparency and encouraging continued data sharing within the scientific community.

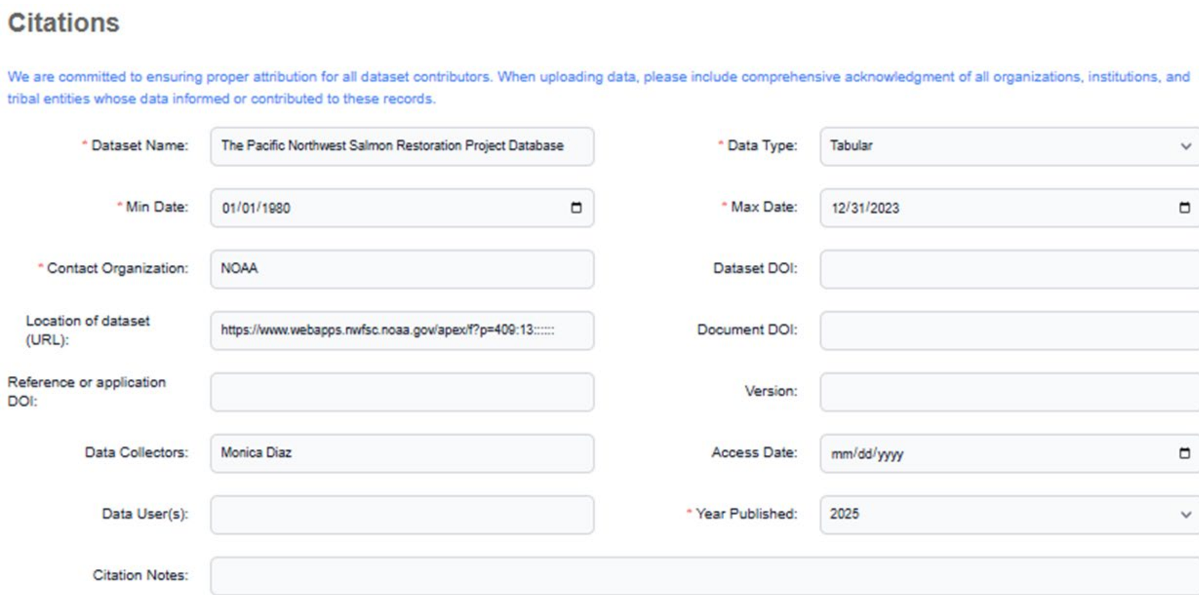


Figure 6: The Citations section of the Data Uploading page

1. *Dataset Name:* This field will autofill from the Title field in Content Description section. However, if you would like to rename your dataset for the purpose of a citation, you can do so here.
2. *Data Type:* This field will autofill from the Title field in Content Description section. This is a drop down menu to describe your file: Tabular, Spatial, Document. For more details regarding this files types, please visit the SDDT Controlled Vocabulary (provide link when ready
3. *Min Date and Max Date:* The start date and end date that pertains to this file being uploaded. This is essential information especially if the data is a time series.
4. *Contact Organization:* Name of the organization or entity that provided the dataset and should be contacted for data use.
5. *Access Date:* If data was downloaded from an existing dataset, provide the date it was downloaded here.
6. *Year Published:* The year the dataset or document was published. If publishing to the SDDT site, then provide the current year.

Contacts:

In this section you will be able to add as many contacts as possible that are associated with the file being uploaded. This section is useful for a user if they have specific questions regarding your file, they will know which contact to reach out to.

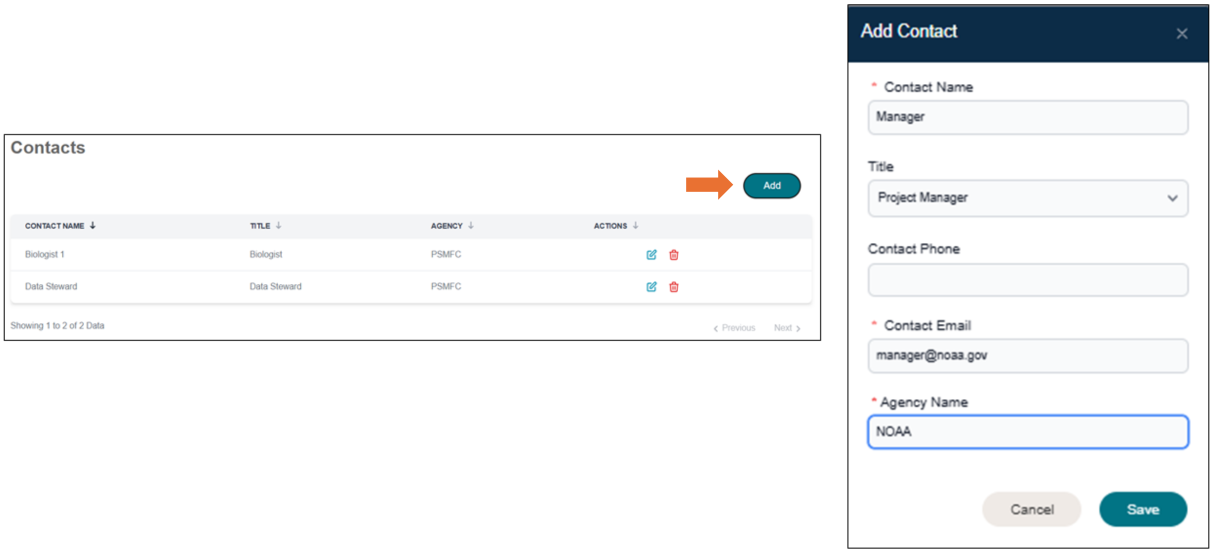


Figure 7: Example of the Contacts section in the Data Uploading page

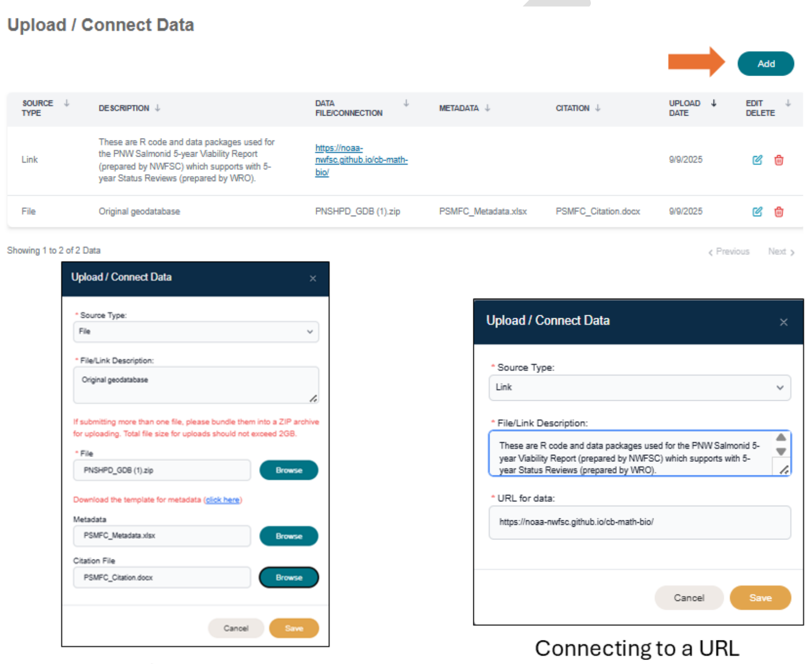
1. Select the Add button on the right side of the Contacts section area. A pop-up screen will appear with the fields needed to add a new contact (see figure 7 above).
2. *Contact Name:* Name of the contact you are adding
3. *Title:* While this is not a required field, it will be important for a user to know the title of the person so they can reach out with questions to the appropriate person.
4. *Contact email*: the current email address of the contact.
5. *Agency Name*: The name of the agency or entity associated with the contact.
6. Select Save when you are finished.
7. The list of contacts will appear in the Contacts section. You are able to edit or delete a contact at any time, even after you have published the data.

Upload/Connect Data

In this section, you will be able to upload the file and or document, as well as the associated metadata and citation document. There is a metadata template provided for you to use if needed, on the SDDT site: https://sddt.psmfc.org/SDDT_Metadata_Template_for_Tabular_Data_2025.xlsx

In this section you can also connect to a URL or DOI where data or to a GIThub link or online dataset.

1. Select the Add button on the right side of this section (see figure 8 below). This will open up a pop-up window.
2. *Source Type:* Select File, if you are uploading a file (must be under 2GB). Select Link if you are connecting to a website (see examples below in figure 8).



File Upload

Connecting to a URL

Figure 8: Example of how to upload a file or connecting to a URL and filling out the fields in the pop-up screen.

3. *File/Link Description:* Briefly describe the file or link you are uploading or connecting to. ***This section is important if you are uploading multiple versions of a dataset. Here is where you can let the user know which version of the data it is.***
4. *File:* Select the Browse button to navigate to the file you want to upload.
5. *Metadata:* If you have a metadata file associated with your data, please upload it here.
6. *Citation file:* A document with the citation you would like users to have if they use your file.
7. Click Save when you are finished.
8. You will see our files in the Upload/ Connect Data section. Here you can edit or delete a file or link at any time, even after it has been published.

Dataset Coverage Area

In this section you can associate your file with a spatial location so that it can be discovered in the Spatial Search. This feature is another way to make your data discoverable on the SDDT site. This feature will take the midpoint of the area you create, and it will be published to the Spatial Search feature.

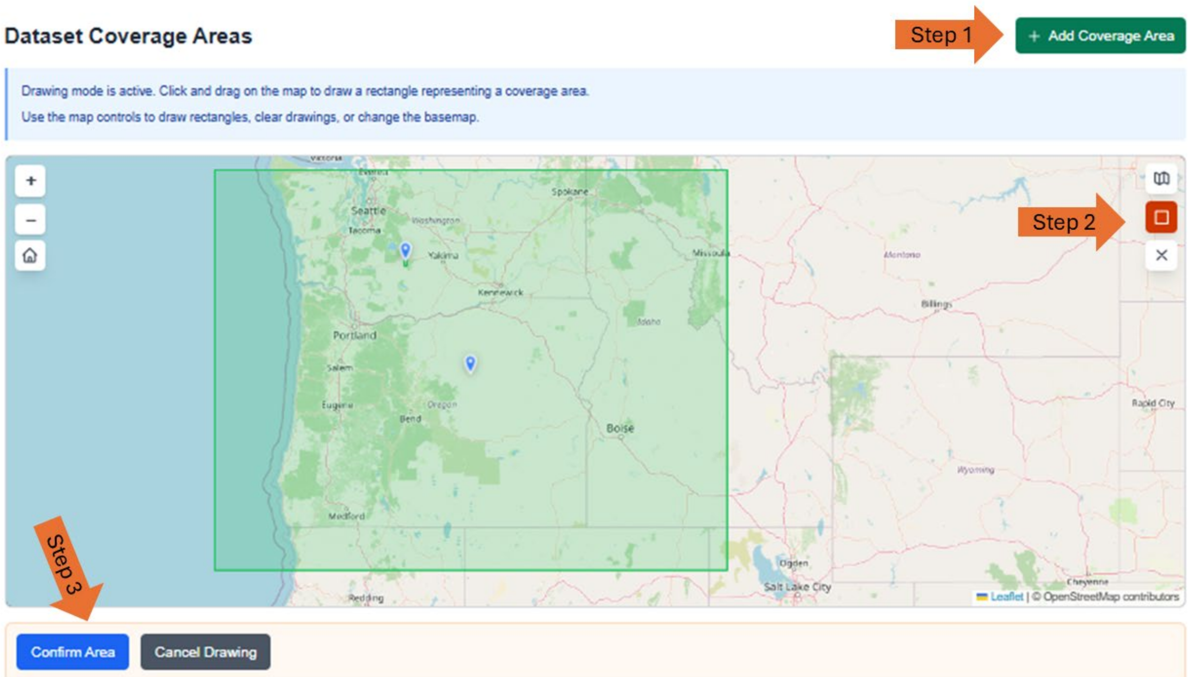


Figure 9: An example of adding a spatial area to a file upload

1. Step 1: Select the green box that says Add Coverage Area, on the right corner of the section (see figure 9 above).
2. Step 2: If the image of a box is not already highlighted orange (in the right corner of the map, see figure 9), select the box to be able to draw a box of the area that best spatially represents the file you are uploading.
3. Step 3: Once you have drawn a box of the area, select the blue box that says Confirm Area on the lower right corner of the section (see figure 9 above). There will be a pop-up window for naming the coverage area (see figure 10).
    a. If you made a mistake, select the X in the upper right corner of the

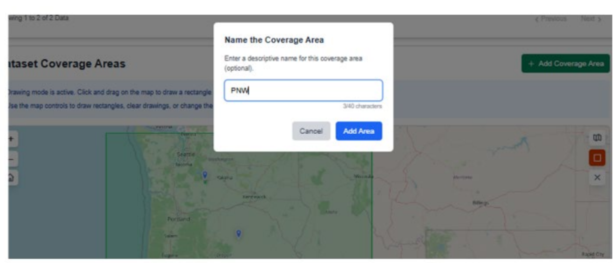


Figure 10: Example of the pop-up window for you to name your coverage area

map, and start again. Alternatively, you can select the gray box that says Cancel Drawing.

4. Multiple spatial locations can be added by repeating this process and naming each location accordingly.
5. Once you are done, you will see all the coverage areas you created on the map in a list below (See figure 11 below). Here you can view or delete a spatial coverage area. At anytime, you can create more coverage areas for your file, even after it has been published.

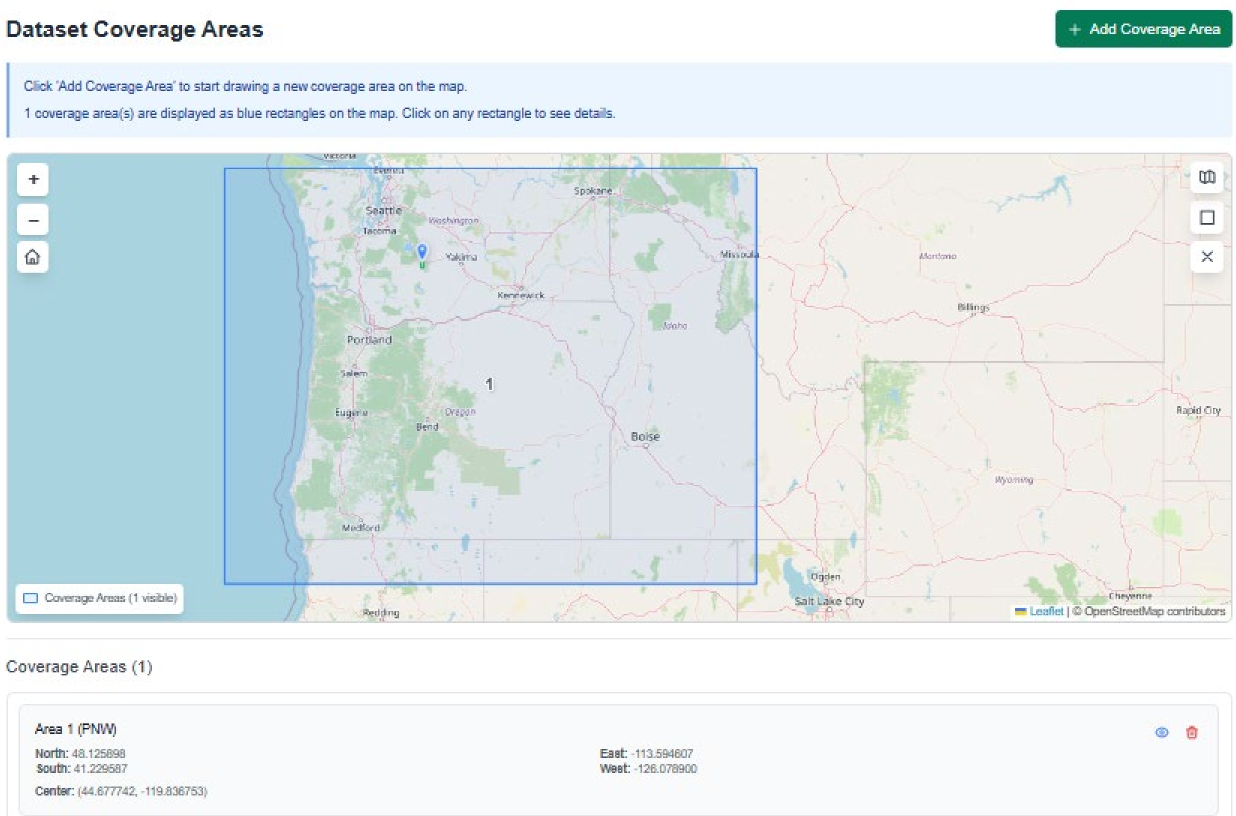


Figure 11: Example of the completed process of adding spatial coverage areas for your file.

Saving your file or link

Once you have all the sections completed and the required fields populated, you can decide to save a DRAFT or PUBLISH your file to the SDDT site.

- *Draft*: If you select Draft, your information will be saved but it is not accessible to the public on the SDDT site. You can also switch the status of your file to Draft, after you Publish your data if you would like to temporarily remove your file from being searchable on the SDDT site.

- *Publish*: When you select Publish, this makes your data accessible to the public and searchable on the SDDT site. When you publish your data, a pop-up window (see figure 12) will open and ask you for a quick description (this can be used for tracking new versions of data) as well as acknowledging the sensitivity of the data.   A note about data sensitivity:
    - Only publish data after you have received permission from all contributing parties
    - Do not publish data that have sensitive location information.
    - **You can only publish Non-sensitive data to SDDT.**
    - **If you have a publish dataset and change the Data Sensitivity to Sensitive, the system will automatically "unplublish" the dataset. Unpublish datasets are not searchable (see figure 12).**

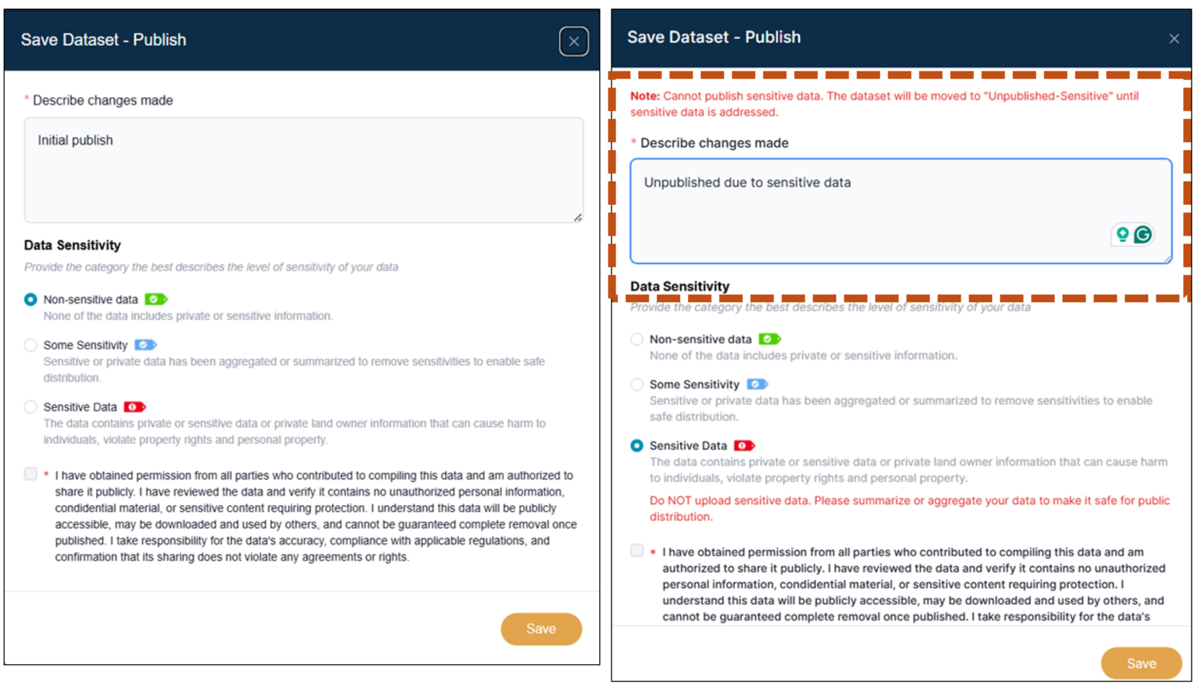
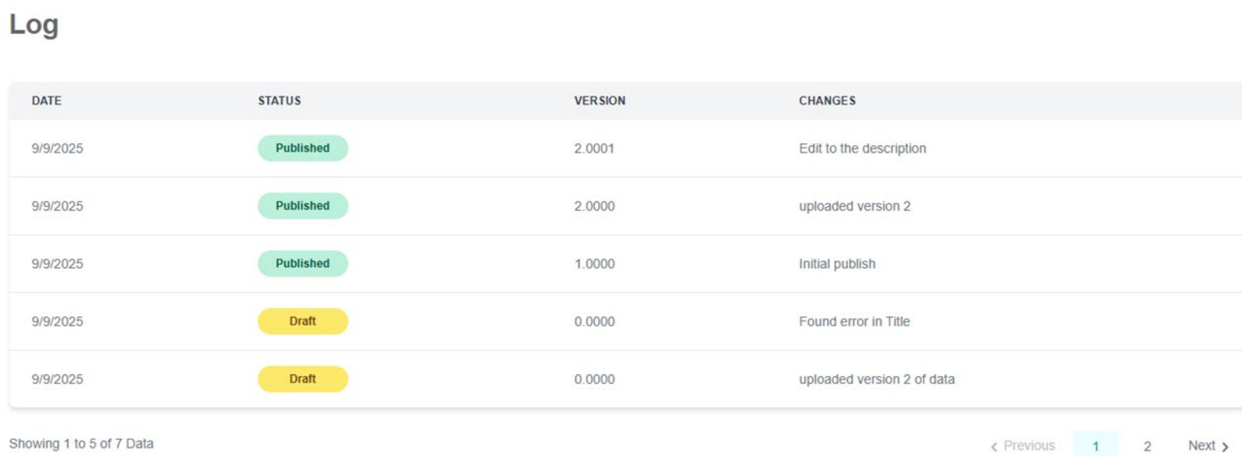


Figure 12: Example of Pop-up window to acknowledge data sensitivity (left) and an example if you change the sensitivity rating after you publish a file (right)

- *Archive*: This will remove your file and its associated information from the SDDT site. It will no longer be searchable. If you archive your file, you will have to go through the process of reloading and filling out the associated information on the Data Uploading page. **Only use this option if you are certain that you would like to remove your data.** To make your data available again you will need to start the process over.

Log File and Versioning

There are two ways that the SDDT site tracks versioning of your file. The site does this so you can track edits that have been made to your data uploading process and if new files are being uploaded under the same Title.



Log

DATE	STATUS	VERSION	CHANGES
9/9/2025	Published	2.0001	Edit to the description
9/9/2025	Published	2.0000	uploaded version 2
9/9/2025	Published	1.0000	Initial publish
9/9/2025	Draft	0.0000	Found error in Title
9/9/2025	Draft	0.0000	uploaded version 2 of data

Showing 1 to 5 of 7 Data ‹ Previous 1 2 Next ›

Figure 13: Example of the log file that tracks edits and version control

- *Minor Version update (decimal .xxxx):* Minor version update automatically occurs when you save edits to the dataset. For example, if you edit the dataset description and save the version will have a minor increment .ooo1 to .ooo2. (see figure 12, first line in the Log).
- *Major Version:* If you have a file that needs to be updated but you want to keep the older file available for download, you can upload a new version of that data. For example if you have a new year of stream temperature data from the same temperature gauge that you have already uploaded (see figure 11, second line in the Log). Uploading a new file will increment the major version X.000) when saving the dataset.
    - You can go to the original dataset, navigate to the Upload/Connect Data section and select Add. Once you have connected to the file to be uploaded, in the File/Link description is where you can type the version of the data and/ or the years associated with the data (see figure 7 in Upload Connect Data section).

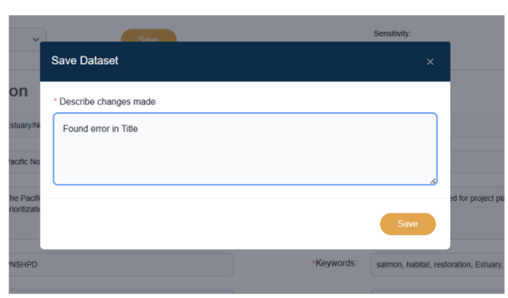


Figure 14: Example of the pop-up window to describe changes to data fields

Archiving file/link

There are two ways to remove your dataset or document from your account.

- In the Data Uploading page, select Archive in the Status drop down menu (See figure 14)
- In your My Datasets page, use the trash icon at the right of the file upload (see figure 15)

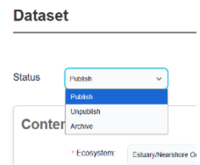
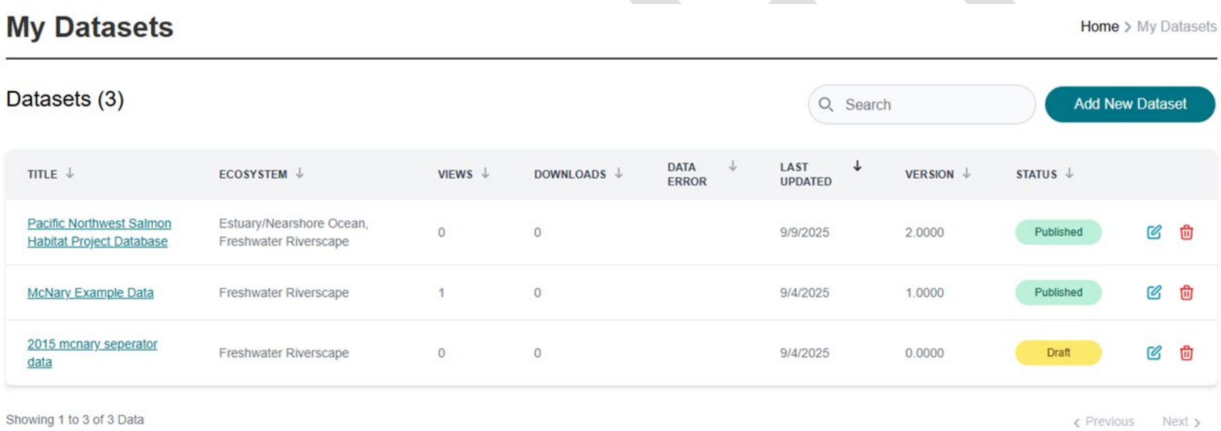


Figure 15: Example of the Status drop down menu

My Datasets page

Once you have finished either saving your file in Draft or Publish Status, you can find all your datasets in your My Datasets homepage (see figure 15 below). Here you can edit or archive your files, at any time. You can also see the status of your files so you know if they have been published or not.



My Datasets

Home > My Datasets

Datasets (3)

Search | Add New Dataset

TITLE	ECOSYSTEM	VIEWS	DOWNLOADS	DATA ERROR	LAST UPDATED	VERSION	STATUS
Pacific Northwest Salmon Habitat Project Database	Estuary/Nearshore Ocean, Freshwater Riverscape	0	0		9/9/2025	2.0000	Published
McNary Example Data	Freshwater Riverscape	1	0		9/4/2025	1.0000	Published
2015 mcnary seperator data	Freshwater Riverscape	0	0		9/4/2025	0.0000	Draft

Showing 1 to 3 of 3 Data

Previous Next

Figure 16: Example of the My Datasets page

Troubleshooting

Common Issues and Solutions

Upload Fails

- Check file size - Ensure it's within 2 GB
- Test internet connection - Ensure stable connectivity
- Clear browser cache - Refresh the page and try again

- Missed Required Fields- Check that all required fields have been filled out.