Data Provider's Guide:
How to Contribute Your Data

Welcome! The Advanced Cooperative Arctic Data and Information Service (ACADIS) is a joint effort by NSIDC, UCAR, and NCAR to provide data discovery, access, and preservation for all projects funded by NSF's Arctic Science Program (ARC).

What is the Purpose of the Data Provider's Guide? The Data Provider's Guide is a reference to help you through the process of submitting data to ACADIS. This guide provides step-by-step instructions for uploading your data and offers recommendations for providing accurate and descriptive metadata.

Who is a Data Provider? Anyone who submits metadata and/or data files to the ACADIS Gateway. You may be a Principal Investigator, Field Scientist, Graduate Student, Co-Principal Investigator, Project Manager, Data Manager, Curator, Librarian, Intern, etc.

Who will assist you in updating or submitting data? The ACADIS Community Support Team is here to assist you with both data and metadata submission needs. You can contact us by emailing support@aoncadis.org or via Skype at *acadis.usersupport*, and a team member will respond within 1 business day. Please include any deadlines for submitting your data. You can additionally request a phone call or Skype session by sending your contact information and availability.

The Data Provider's Guide contains the following sections:

A. Structure of the ACADIS Data Collection

Data and metadata collections within the ACADIS Data Gateway are organized by NSF projects. Each project will have one or more associated datasets, and each dataset will contain one or more data files. Figure 1 explains the relationship between projects, datasets and nested datasets, and the associated metadata.

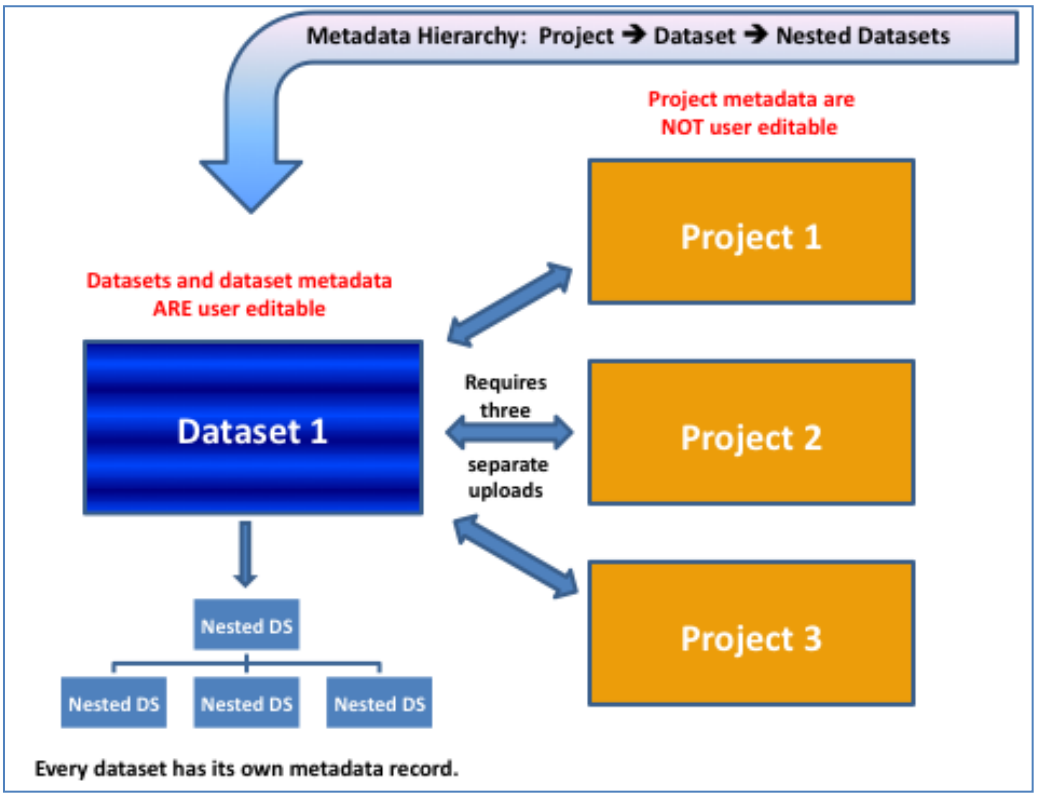

Figure 1: Structure of the ACADIS Data Collection

In Figure 1, data files are not shown. Each dataset, including nested datasets, will have data files in a directory structure determined by the data provider. Some of the elements of the ACADIS data collection are described below.

Project: ACADIS staff design project and dataset hierarchies based on the NSF Award Information available and the PI or data provider's current and foreseen structure of their datasets. From this information, we determine 'What is a Project?' based on the needs of our data providers. Project-level metadata are pre-populated in the ACADIS metadata database by ACADIS Community Support staff. If you would like to submit data to a project not current represented in the ACADIS Gateway, please contact ACADIS Community Support and request that a new project be created. Please include your award number in this request. ACADIS Community Support will respond by acknowledging that the project has been created and that you can proceed with data submission. If

there are errors in the project information, or if you would like to contribute to or clarify the project metadata created by ACADIS Staff, please contact Community Support with your feedback.

Dataset(s): As a NSF OPP ARC Data Provider, you are responsible for creating datasets in the ACADIS Gateway; this process includes both entering metadata and uploading data files. It is up to the Data Provider to decide what collections of data files will constitute individual datasets and how these should be organized in the ACADIS Gateway. What constitutes a dataset can be somewhat arbitrary. A dataset can be a collection of data for a specified time period, for a specific parameter, or for a specific location. Datasets can also be nested under other datasets. ACADIS Community Support is happy to provide insight on how to organize or nest your data.

Data file(s): Each dataset can consist of one, several, or many data files depending on how you, the Data Provider, prefer to organize your data. It is recommended that a ReadMe.txt file, or [dataset_title}_documentation.txt be included in the file structure along with your data files. Please see Appendix B for more information about contributing additional documentation to be associated with your datasets. This documentation allows you to be more descriptive and provide additional information others may need to reuse your data properly.

B. Registration and Login

Note that registration with the ACADIS Gateway is only required if you would like to 1) submit metadata or data files, or 2) edit existing dataset metadata. Registration is not required in order to search for and download data. To register with the ACADIS Gateway, please follow the instructions below:

1. Open the ACADIS Gateway in your browser. We recommend using Firefox, Safari, or Google Chrome in that order.
2. Scroll to the upper right hand corner of the website. If you have not registered before, click 'Request an Account' (see Steps 3-6). If you are returning to ACADIS to edit metadata or submit new datasets, click 'Login' (see Step 7).
3. Fill in the required registration information. Please use a current email address that you frequently monitor. Also, please choose a unique user name that is easy to remember.
4. After you have submitted your registration information, you will receive an email indicating that your account has been created. Please follow the link in the email to confirm your account.
5. You should receive a second email including your account confirmation and a request to email ACADIS Community Support with your award number.
6. After emailing ACADIS Community Support with the requested information, we will set up your project and make sure you have permission to upload. We will send you a final email with instructions on uploading your files.
7. Login with your user name and password. If login was successful, you should land on the home page. If you forgot your password, please click 'Forgot your password' below the user name and password boxes on the Login screen in order to change your password.

C. Creating a Dataset

If you would like to add a new data directly related to a project, please see Steps 1-2. If you would like to nest a new dataset under an existing dataset, please see Step 3. For additional information on submitting your data files, please see Section E: 'Submitting and Replacing Data.' For clarification, please contact ACADIS Community Support;

1. After logging in, click the 'Contribute [Meta]data' tab on the horizontal navigation bar under the ACADIS logo. When you are logged in, a page entitled 'Editable projects for: [your name]' will display a list of projects you can edit. From this list, choose the project under which you would like to submit your data. The project title corresponds with the one you provided to ACADIS Community Support Staff or that was referenced from your award number.
2. Clicking on a project title will open the metadata entry form for submitting dataset metadata. Please reference Section D: 'Contributing and Saving Metadata' for instructions on how to edit this form. Once you have completed at least the required fields, click 'Save Metadata' at the bottom of the form to store the metadata you entered. You will be presented with the new catalog page for your dataset where you can review the entered metadata.
3. If you would like to nest a dataset under existing datasets, search for your Project or dataset name through the Search box in the top right area of the screen, or use the horizontal navigation bar links to browse through PI or Project names. When you have located the appropriate project, scroll down to the 'Nested Datasets' section and click on the title of the dataset to which you would like to contribute a nested dataset
4. On the catalog page for that dataset, click on the 'Contribute [meta]data' tab and then on 'Create new Dataset: Enter Metadata.' This opens the metadata entry form for submitting dataset metadata. Please reference Section D: 'Contributing and Saving Metadata' for instructions on how to edit this form. Once you have completed at least the required fields, click 'Save Metadata' at the bottom of the form to store the metadata you entered. You will be presented with the new catalog page for your dataset where you can review the entered metadata.

D. Contributing and Editing Metadata

Descriptive metadata is required for each dataset submitted. You must complete the metadata entry form whether you will be uploading data files or linking to data files stored elsewhere.

There are two ways to reach the metadata entry form:

1. Click on the 'Contribute [meta]data' menu selection. This should open a page listing the projects for which you can create or edit dataset metadata.
2. Navigate to your existing dataset using the browse tabs or the search box. Click on the 'Contribute [meta]data' tab at the top of your dataset catalog page. From the list of options, click on either 'Create new Dataset: Enter Metadata' or 'Edit this dataset metadata.'

You as a Data Provider MUST PROVIDE a minimum set of required metadata before uploading data file(s) to ACADIS. The required metadata fields are listed first on the metadata entry form and are indicated with an asterisk (*). Please follow these best practices for contributing metadata:

1. *Save Frequently:* Any entries and/or changes you make will only be saved if you click on the 'Save Metadata' button at the bottom of the form.
2. *Ease of Entry*: We recommend completing the metadata form with limited required metadata entered and then saving your entries. This allows to you quickly save the record and then return to the record to casually edit the existing metadata. This prevents your session from timing out or clearing if there is an error with your submission. Return to edit the metadata page by selecting 'Contribute [meta]data' on the newly created dataset catalog page and then choosing 'Edit this dataset metadata.'
3. *Complete and Accurate*: Please note that it is important to fill in the metadata fields with information as accurate and complete as possible. The metadata you provide are used by several functions, including search and browse. Invalid or inaccurate metadata may prevent your datasets from being properly accessed and displayed. When in doubt about how to complete a metadata field, contact ACADIS Community Support.
4. *Follow an Example*: See 'ACADIS Metadata Record Example' in Appendix A to reference a sample complete metadata record. See the following 'Metadata Fields and Definitions' section for recommendations on each specific field.
5. *Provide Additional Documentation:* It is also suggested that each dataset include a 'readme' file, or dataset documentation, with any and all additional information needed to properly use the data files. Recommended contents are in Appendix B.
6. *Understand the Metadata you are Selecting or Creating:* Be aware of the field definitions when you are filling out metadata. Submitting consist metadata to current standard definitions is important in order to achieve interoperable and well-described metadata. ACADIS currently subscribes to the NASA Global Change Master Directory's (GCMD) vocabularies and incorporates ISO Topics in order to be compliant with GCMD and ISO science metadata standards. ACADIS metadata are compliant with GCMD Directory Interchange Format (DIF), the International Polar Year metadata profile, and THREDDS.
7. *Do not include phone numbers or email address in the metadata:* Please be careful that you avoid adding any contact information for any other personnel involved in the project, especially in the dataset summary. We suggest this in order to prevent private information from being captured by online bots and other intruders. If you need someone's contact information, we suggest searching through his or her workplace's person directory or contacting us directly.

Metadata Fields and Definitions:

Note that any fields with asterisks (*) are required. Any comments that are Italicized are those you should keep in mind when completing the metadata entry form.

- ***Project**: The project shown should be the project that funds your data collection, or that your data has been collected to support. This field is pre-populated. Please contact ACADIS Community Support if there is an error in the project name, if your project is not listed, or if you have any further questions.
- ***Title**: This is the title of the dataset. Please make this title unique and descriptive of the data content. Put the most important words first and do not use only dates for titles. Make sure that the dataset title is consistent with your other documentation. If the title doesn't fit in the entry window it is probably too long. *Keep in mind that, like the title of a research paper, the title of a dataset may be used in a citation.*

- ***Short Name:** A one-word label for the dataset. This could be an acronym or abbreviation of the dataset title and must be unique.
- ***Description**: This is a paragraph that describes the 'who, what, where, when, and why' of the dataset you are submitting. Remember, this text should describe the specific dataset, not the project as a whole. *Beware that copying/pasting from some documents may result in embedded control characters which may or may not be supported by the ACADIS Gateway*. *Also, consider referring to published documents when more detail is required to describe a dataset adequately. Be careful to exclude email addresses and other contact information in the dataset summary.*
- ***Author**: The people or organizations responsible for the intellectual work to develop the dataset, as should appear in a citation. Data authors should be listed in order of pertinence to the dataset. If there is no pertinence, order alphabetically by last name. ACADIS recommends using initials for representing first and middle given names. Below is an example of data authors for a dataset:

  Conover, A., R. Greene, T. Johnson, M. Richardson, and P. Smith.
- ***Location Keyword(s)**: Choose one or more location names that describe where your data were collected.
- ***Platform Keyword(s)**: Choose the platform or platforms for the instruments that acquired these data. Please contact ACADIS Community Support if your data cannot be accurately described by the contents of this field.
- ***Instrument Name(s):** Choose one or more instruments from the list. If the required instrument is not in the list, select 'other' or please contact ACADIS Community Support to request additional instrument names.
- ***Science Keyword(s)**: Choose one or more from the NASA Global Change Master Directory (GCMD) terms.
- ***ISO Topic(s)**: Please select at least one ISO Topic from this list. ISO Topic is included in the ACADIS metadata profile to be compliant with international metadata standards.
- ***Metadata Contact**: The Metadata Contact is responsible for the content of the metadata record. This person is usually the same as the Data Provider. As a Data Provider, you will be added to this list when you are given permission to edit datasets. Please contact ACADIS Community Support if a different name is needed. If the responsibility shifts from the original author to another person (*e.g*. from an investigator to a data manager), this field should be updated to the newly responsibly person.
- ***Data Center Contact**: The person (or data center) responsible for the distribution of the data. If data are submitted to ACADIS, please select the ACADIS Community Support Team member who initially responded to your request. If data are archived elsewhere and only metadata are submitted to ACADIS, select the person responsible for the data. If the appropriate name is not listed, please contact ACADIS Community Support.
- ***Distribution Format(s)**: Choose the file format(s) of your data.
- **Begin and End date**: Choose from the drop down menus to provide the temporal coverage of the dataset. If hour and day of month are known, please include them.
- **Southernmost and Northernmost Latitude/Westernmost and Easternmost Longitude**: We highly recommend that you supply the values for spatial coverage. Enter this metadata as you would portray a box around your dataset in term of minimum and maximum latitude and longitude. If the data are at one geographical location only, enter the latitude and

longitude of that single point as both extents. For a moving platform, use a box that includes the track of the platform. Position to three decimals is usually sufficient. Although not required, this field is highly recommended. The following guidelines from the GCMD may be useful;

  - **Southernmost_Latitude (Minimum Latitude)**: The southernmost geographic latitude covered by the data. From: 0 to 90 deg for northern latitude or 0 to -90 deg for southern latitude. For example, 60 will be 60 degrees north, -60 will be 60 degrees south.
  - **Northernmost_Latitude (Maximum Latitude)**: The northernmost geographic latitude covered by the data. From: 0 to 90 deg for northern latitude or 0 to -90 deg for southern latitude. For example, 60 will be 60 degrees north, -60 will be 60 degrees south.
  - **Westernmost_Longitude (Minimum Longitude)**: The westernmost geographic longitude covered by the data. From: 0 to 180 deg or 0 to -180 deg. The Prime Meridian (PM) is 0 degrees, measured positive (+) eastwards of the PM and negative (-) westward of the PM. For example, 45 will be 45 degrees east and -45 will be 45 degrees west.
  - **Easternmost_Longitude (Maximum Longitude)**: The easternmost geographic longitude covered by the data. From: 0 to 180 deg or 0 to -180 deg. The Prime Meridian (PM) is 0 degrees, measured positive (+) eastwards of the PM and negative (-) westward of the PM. For example, 45 will be 45 degrees east and -45 will be 45 degrees west.

- **Frequency(ies)**: Choose one or more options from the list that best describe your dataset. Choose the frequency(ies) that best approximate(s) the reporting frequency for your data.
- **Spatial Type(s)**: Choose the data type(s) that most closely match(es) your data.
- **Resolution(s)**: Choose from the spatial resolution ranges given by the GCMD keywords. For some datasets (gridded satellite image data, for example), resolution can be harder to define. Choose the resolution(s) that best describe(s) the data. You can include more detailed information in the additional documentation or readme files you provide.
- **Progress**: This is meant to describe the state of the data being submitted. 'Planned' refers to datasets to be collected in the future and are thus unavailable at the present time. 'In Work' means the data are preliminary or data collection is ongoing. 'Completed' refers to a dataset for which no updates or further data collection will be made.
- **Data Access**: This is a URL to data *not* submitted to the ACADIS archive. *If you submit data to ACADIS, leave this blank.* For external data, this will be a link to the data location and metadata.
- **Related Resource:** These are references (citations) to work that use the data, link(s) to full product documentation, and/or links to related data. Insert a URL or citation and then tag it with a 'Purpose' from the pull-down menu.
- **Dataset Language**: If the dataset (data and/or documentation) is in a language other than English, please enter the language.
- **Access Restrictions**: Datasets submitted to ACADIS are unrestricted in access as soon as they are uploaded. If you do not want your data to be openly accessible, you must speak with your NSF manager about exceptions. Exceptions may include any special restrictions,

legal prerequisites, limitations, and/or warnings on obtaining the dataset. Additional information and examples are available from the GCMD.

- **Use Constraints**: Describe here how the data may or may not be used (after access is granted) to assure the protection of privacy or intellectual property. This includes any special restrictions, legal prerequisites, terms and conditions, and/or limitations on using the dataset. Data providers may request acknowledgement of the data from users and claim no responsibility for quality and completeness of data. Additional information is available from the GCMD.

Click on the 'Save Metadata' button when you are finished editing the metadata. Edits to the metadata will not be saved unless you click this button.

E. Submitting and Removing Data

You can submit data files when you create a dataset or update data files in an existing dataset.

Requirements:

Individual file size is currently limited to 500 MB.

File names can include upper- and lower-case letters, numbers, spaces, dashes, and underscores. Do not use special characters including "/ \ : * ? < > [ ] & $ | % !.

To submit data files to a new dataset:

From the metadata record page for your dataset, click on the 'Contribute [meta]data' tab. You can upload up to five files at once with "Basic File Upload", or as many as you like from a selected directory using "Bulk File Upload" (with an HTML5-enabled browser such as Internet Explorer 10.0+, Safari 6.0+, Firefox 3.6+, Chrome 3.0+, Opera 4.0+).

To remove files for an existing dataset:

Contact ACADIS Community Support to request the deprecated files be deleted.

For help uploading or if you have trouble:

Contact ACADIS Community Support (support@aoncadis.org).

Appendix A: ACADIS Metadata Record Examples

Project level information (project metadata) is illustrated in Figure 2. Figure 3 shows the metadata for a dataset that falls under the project from Figure 2.

Figure 2. Project level information (project metadata). ACADIS support staff create this metadata based on NSF grant information.

Figure 3. A dataset metadata example.

Appendix B: Creating and Submitting Data Documentation

The 'readme' file is a critical part of dataset documentation. It should contain enough information for a researcher who is unfamiliar with the project under which the data were acquired to use the dataset. It can begin with the same summary paragraph that is used in the metadata. Following are headings and a suggested outline for dataset documentation.

Readme.txt

Dataset Title [This is also a metadata field.]
This field helps others find your data.

Please give your dataset a descriptive title that is less than 220 characters. It should be descriptive enough so that when a user is presented with a list of titles the general content of the dataset can be determined. For example, *Aerosols* would not be an adequate dataset title, but *Aerosol characterization and snow chemistry at Terra Nova Bay* would. So, if it can be done without making the title too long, include parameters measured, geographic location, instrument, investigator, project, and temporal coverage.

Examples:

- National Solar Radiation Data Base Hourly Solar Data from Alaska, 1961-1990
- Comprehensive Ocean - Atmosphere Dataset (COADS) LMRF Arctic Subset
- Daily Precipitation Sums at Coastal and Island Russian Arctic Stations, 1940-1990

Summary Description [This is also a metadata field]
An 'above the fold' overview of everything someone might need to know to decide if the dataset is something they can use., about ½ page long at maximum.

This is a paragraph that describes the 'who, what, where, when, and why' of the dataset you are submitting with this metadata. Also, consider referring to published documents or web sites when more detail is required to describe a dataset adequately.

Contacts
Any relevant people, with their titles, and role.

Background
Any contextual information, for example, is the dataset part of a larger experiment or collection? Were certain aspects of the data acquisition procedure notable or unusual?

Detailed Data Description
Here are some example subcategories. Use what seems logical, putting yourself in the place of a researcher outside your field who would like to use your data.

- Parameters

- Data File Format
- Sample Data Record
- File Naming Convention
- File Size
- Spatial and Temporal Coverage and Resolution
- Quality Assessment

Data Acquisition and Processing

Be as descriptive as possible, with references to instrument manuals, standards, or other works where applicable. For example:

> Snow depth on sea ice measurements were acquired according to the method detailed in Chapter 3.1 (Strum, 2009) of *Field Techniques for Sea Ice Research* (Eicken, 2009)'. Another example: Sea ice optics measurement sites were selected following the guidelines set forth in Chapter 3.6, Section 3.6.3, *Methods and Protocols* (Perovich, 2009) of *Field Techniques for Sea Ice Research* (Eicken, et al., eds., 2009).

References and Related Publications

These can include papers that use these data or like data. References that describe how measurements are taken are especially valuable. For example:

Strum, M., 2009. Field Techniques for Snow Observations on Sea Ice, in *Field Techniques for Sea-Ice Research*, Edited by H. Eicken et al., University of Alaska Press, Fairbanks, 588pp; ISBN 978-1-6022230-59-0

Perovich, D. 2009. Sea Ice Optics Measurement , in *Field Techniques for Sea-Ice Research*, Edited by H. Eicken et al., University of Alaska Press, Fairbanks, 588pp; ISBN 978-1-6022230-59-0

H. Eicken, R. Gradinger, M. Salganek, K. Shirasawa, D. Perovich And M. Leppa¨ Ranta, *eds*. 2009. *Field techniques for sea ice research.* Fairbanks, AK, University of Alaska Press. 588pp. ISBN-10: 1-602230-59-5, ISBN-13: 978-1-602-23059-0,

Acknowledgments

As a rule, include the grant number or numbers and funding agencies that supported the work.

Document Information

Here name the document author(s), date and the date it was created or revised. Including the date is important.

If a revision was made, say when it was made and briefly, what the nature of the revision was. For example, 'This readme documentation file was revised on 4 Feb 2011 to add information about a new instrument used to measure snow density, and about the resulting new additional data files.'