

[LTA Home](#)[IGLOO Home](#)[NSIDC Home](#)

---

# NSIDC Data Management Policies, Version 2.0, February 2004

compiled by Teresa Mullins and Ruth Duerr

approved by the members of the LTA team July 14, 2003

approved in principle by the NSIDC Management Team July 15, 2003

Point of contact: Clark Judy, Deputy Director

## Revision History

Version 1, July 15, 2003 - approved by NSIDC Management Team

Version 2, February 11, 2004 - incorporates edits by IT team reflecting new NSIDC organizational structure

## Table of Contents

- [1.0 Introduction](#)
- [2.0 Data Solicitation and Acceptance](#)
- [3.0 Levels of Service](#)
- [4.0 Metadata and Data Format Standards](#)
- [5.0 Data Set Documentation](#)
- [6.0 Archive Policies](#)
- [7.0 Architecture and Security](#)
- [8.0 Data Tools](#)
- [9.0 Data Deletion and Retirement](#)
- [10.0 Data Access](#)
- [11.0 References](#)
- [Appendices](#)

## 1.0 Introduction

### 1.1 Purpose and scope:

The National Snow and Ice Data Center (NSIDC) Data Management Policies document outlines the Center's policies for acquiring, accessing, managing, distributing, and archiving data and metadata. These policies follow traditional archival standards, and reflect NSIDC's mission statement:

*"NSIDC/WDC will make fundamental contributions to cryospheric science and will excel in managing*

*data and disseminating information in order to advance understanding of the Earth system."*

NSIDC's mission statement lists priorities of data management, information dissemination, and cryospheric research, with the understanding that these activities support and further the others. Reputable and consistent long-term data enhance scientific understanding of the cryosphere and its impact on and response to global climate. NSIDC ensures quality and reliability by taking a systematic approach to scientific stewardship of its data.

NSIDC is part of the Cooperative Institute for Research in Environmental Sciences at the University of Colorado, Boulder, and is funded at the project level by national agencies. Primary among these agencies are the National Aeronautics and Space Administration (NASA), the National Science Foundation (NSF), and the National Oceanic and Atmospheric Administration (NOAA). NSIDC began as the World Data Center for Glaciology, Boulder, and continues in this capacity. Because NSIDC is a matrixed organization of various programs and projects, the Center must develop data management policies that are stringent enough to guarantee best practices, adherence to standards, and strong and consistent levels of service, yet are flexible enough to adapt to changing programmatic relationships and needs.

NSIDC has established a collection scope as follows:

*"NSIDC will provide long-term data management, including data archival, as an essential part of its mission,"*

and

*"NSIDC will archive both cryospheric data and data from programs or instruments deemed important to the cryospheric community."*

NSIDC programs adhere to the collection scope in the data they accept, disseminate and archive. By focusing the subject matter covered, NSIDC better serves current and future data users by providing quality levels of service, including knowledgeable support and creation of applicable tools and guidance documents.

Typically, archives are defined as the non-current but still useful records preserved by an organization or institution, or the repository itself. The term archives can also refer to organizations or institutions that collect and acquire records or data from outside sources, usually to serve researchers outside that organization.<sup>[1]</sup> Because NSIDC functions as both a long-term archive and a short-term or "active archive," its data management roles and responsibilities are larger than a typical archive. NSIDC bases its data management policies for all data on proven archival principles and best practices. These principles are:

- Archival records exist to be used, not merely saved for their own sake.
- Some records ought to be preserved long-term, even after their immediate usefulness has passed.
- Archival records ought to be preserved as completely and coherently as possible, including critical information about context and connections.
- Archival records ought to be organized properly and in a timely way so they can be used.
- Sensitive information and information given in situations presumed to be private should be protected from use as long as that sensitivity remains.
- Archivists should administer their collections equitably and impartially.
- Archival repositories ought to cooperate in preserving historical records.<sup>[2]</sup>

NSIDC incorporates these archival principles into its data management processes and they, along with

NSIDC's mission statement and collection scope, serve as the core to NSIDC's Data Management Policies. Specifics of these policies will be elaborated in the remainder of this document.

## 1.2 Document Format

Throughout this document ***bold and italic*** type is used to highlight policy statements.

## 1.3 Applicability

***The NSIDC Data Management Policy will be applied to all programs, projects, data sets and related data management activities that are the responsibility of NSIDC. A data set is the responsibility of NSIDC if NSIDC is the primary or backup archive for it. The policies in this document are effective as of July 15, 2003. Efforts will be made to bring data sets and tools accepted prior to July 15, 2003 in compliance as funding and resources permit.***

## 1.4 Change Control

NSIDC recognizes that to keep this document pertinent and effective, it should be reviewed frequently. ***The Deputy Director (DD) is responsible for the maintenance of this plan. The Program Coordination Board may make periodic minor revisions or clarifications with the approval of the DD. The plan should be reviewed comprehensively on a regular basis (e.g., annually) by representatives from across all areas of the Center (Science, Service Groups, Programs), perhaps during a retreat.***

## 1.5 Document Structure

This policy statement consists of sections that touch on all aspects of NSIDC data management. This document is not meant to serve as a manual of data management processes. Nonetheless, policies are presented in a chronological fashion, from data acceptance through dissemination to users.

Data acquisition issues are presented under [Data Solicitation and Acceptance](#) policies and cover the various means by which data may be identified and acquired. Closely associated with this is data deletion and retirement and the data center policies for decommissioning data sets.

Components of [Levels of Service](#) follow with discussions of NSIDC's recognition that not all data sets can or should receive the same levels of stewardship. Minimum service level criteria are described and preferred service levels are described as appropriate.

Policies for cataloging and archival processes that allow access to and sharing of data are covered under [Metadata Standards](#). NSIDC attempts to incorporate and use all appropriate national and international standards, and to adapt to changing standards. This section also discusses standards for documentation and for electronic and media formats that are accepted by NSIDC.

The section on [Archive](#) policies, including media management, examines data handling and NSIDC's contingency strategies to ensure long-term data retention. Closely associated with this are [Architecture and Security](#) that outline usage of NSIDC technology platforms.

Because specific tools are needed to optimally use some data, the section on [Data Tools](#) discusses those currently available at NSIDC, and the policies for the creation and support of new tools.

Since NSIDC serves as both an active and long-term archive, and because program relationships may

fluctuate over time, [Data Retirement and Deletion](#) policies are spelled out to ensure minimum disruption to users and data providers.

Finally, [Data Access](#) presents policies that allow equitable access to data in accordance with federal laws and NSIDC's mission.

Program specific data management policies are outlined in the appendices.

## 2.0 Data Solicitation and Acceptance Policies

Data become the responsibility of NSIDC through a variety of official program agreements, formal and informal PI relationships, data rescue projects, and data exchanges. Thus, data can be both solicited and tendered, can arrive with or without funding for publication and archival, and can have various degrees of value to current and future user communities. New additions to the NSIDC catalog/collection begin with identifying the data to be added, followed by appraisal and acquisition phases. *All NSIDC programs recognize [the collection scope of the Center](#) and acquire data within that scope. At the same time, NSIDC acknowledges the individual data management requirements of the programs/projects it supports, and works to meet these needs.*

### 2.1 Data Solicitation Policies

*NSIDC and the programs that make up NSIDC actively solicit new data sets for archival and distribution.* NSIDC's collection scope to "archive both cryospheric data and data from programs or instruments deemed of importance to the cryospheric community," requires NSIDC to be an active collector of new data and data sets that are gathered or discovered. Submission forms for PIs and other repositories of data are available online on the NSIDC web page, as well as on individual program pages. Many programs for which NSIDC is the official data archive are required to solicit data from data providers as part of their grant funding directives. NSIDC Programs with active solicitation policies include the International Arctic Research Center (IARC) Frozen Ground Data Center (FGDC) and Global Geocryological Data (GGD) System, the NSF Antarctic Glaciological Data Center (AGDC), and the NSF Arctic Systems Science (ARCSS) Data Coordination Center. For more information about the specific data solicitation policies of these programs, please refer to the appendix. For some programs, NSIDC will not archive all of the data, but will archive only the metadata that describes the data and gives the data location.

### 2.2 Data Acceptance Policies

Data submission forms encourage the submission of data to NSIDC, while informing data providers and repositories that the submission serves only to begin a discussion of the appropriateness of the material for NSIDC, timeline for publication and archival, and the costs involved. When a data submission form has been received by NSIDC for a data set not covered by any specific Program, *the Program Coordination Board (PCB) with relevant User Services and scientific input decides whether to accept the data and assigns a point of contact for the data.* Acceptance policies are outlined below.

*All accepted data include a Deed of Gift or other documentation that spells out the rights and responsibilities of NSIDC, the funding program, and the data provider.* An example of a Deed of Gift statement is provided in [Appendix A](#). *Information that must be provided includes: Name and contact information of data provider, restrictions on the use of data, such as copyright restrictions/hold until published/restrictions on use outside of the United States, etc., funding provided for the archiving of the*

*data (if any), terms of dispersal should the data be retired, deleted from the archive or updated [see [Data Deletion Policies](#) section]. Additional information that should be solicited includes context of the data within the broader cryospheric environment and provenance of the data. Whenever possible, data are accepted only in specific acceptable formats that are outlined in the [Data Format](#) section of this document. If it is not possible to receive the data in one of these formats, the designated point of contact will seek funding to migrate the data to an acceptable format.*

### 2.2.1 Program Data and Agency Requests

Each program at NSIDC has an individual whose responsibilities include the collection, verification, archival and publication of new data. ***Program Managers or their representatives follow both NSIDC policies and their program's policies for data management (see appendices for more information.)***

NSIDC serves as the sole archive of cryospheric data, and of data of importance to the cryospheric community, for many agencies. Data are occasionally submitted from, or at the request of, one of these agencies. ***In many cases, the rights and responsibilities of NSIDC and the funding program are spelled out in the contract or grant. In other cases it will be necessary to pursue a separate agreement with the funding program to provide a blanket "Deed of Gift" or equivalent document covering all relevant data sets. In these cases, the Program solicits name and contact information as well as data set context and provenance information from the data provider, provides information about NSIDC's implied rights and responsibilities to the data provider, and solicits agreement or disapproval of these rights and responsibilities from the data provider. If the data provider does not agree to the implied rights and responsibilities, the program must pursue an individual data set Deed of Gift in order to accept the data.***

### 2.2.2 PI-Tendered Data Management Requests

Many data sets come to NSIDC as a result of PI offerings that are unrelated to any program. ***The ~~NMT~~Program Coordination Board's informed (USO, Science) appraisal of these data determines their appropriateness within NSIDC and the appropriate program to which they should be affiliated. NSIDC prefers that all data sets come to NSIDC with identified funding for their continuation and management.***

***NSIDC recognizes that occasionally data of value to the cryospheric community become available for archiving, though without needed funding or support. When possible, the PCB or designated lead Program works with the data provider to identify possible future funding opportunities, such as data rescue or exchange. If the PCB or designated lead Program feels that the value of the data is great enough, the data are accepted as "orphan" and given a lower level of service to support it. (Please see [Levels of Service](#).) The PCB or designated lead Program will continue to examine future funding opportunities to ensure the proper long-term management of the data at a full level of service.***

***In the appraisal of new, PI-tendered data, the PCB considers the following:***

- ***appropriateness to NSIDC's [mission statement](#) and [collection scope](#);***
- ***its relationship to other data archived at NSIDC and thus its ability to illuminate or increase the value of existing data;***
- ***the cost of archiving the data (including space and resource needs);***
- ***the value in strengthening bonds with funding agents and other agencies;***
- ***outreach opportunities that fit within the NSIDC Outreach goals and objectives.***

### 2.2.3 Targets of Opportunity

Individuals associated with NSIDC, in the course of their research, data management, and data dissemination work, become aware of data deemed of value to NSIDC. Frequently these data are the results of current or planned research projects or fieldwork. *NSIDC works with the individuals involved to determine the appropriateness of the forthcoming data, and the DD makes sure that proper data management policies are involved in the transfer of these data to NSIDC.*

### 2.2.4 Data Exchanges

On occasion, NSIDC receives data from a data provider in exchange for NSIDC data products, to encourage international cooperation and dissemination of NSIDC data. *The PCB's appraisal of these data determines their appropriateness within NSIDC and the appropriate program to which they should be affiliated.* It is recognized that these data seldom have available funding for archival processes.

### 2.2.5 Other

Some data sets and data products within the collection scope of the Center are created in-house by NSIDC personnel. *The acceptance of these data sets requires the same process as does the acceptance of new data, and steps are taken to ensure a high quality product that maintains its ties, through metadata and documentation, to the original or raw data product.*

## 3.0 Levels of Service

*NSIDC supports a range of services, broadly categorized as providing service to the user or for the data.* Categories of services for the user include user services, data set software development, data set documentation, and data distribution. Categories of services for the data include data ingest, data processing, and data archiving. To aid in discussions with potential data providers, levels of services for each of these categories are described in [Appendix B](#). In each category, the levels of service are described in order of increasing cost. In addition, depending on the needs of the data provider or user community, a wide range of value-added products or services may be developed as warranted.

## 4.0 Metadata and Data Format Standards

In keeping with data management best practices, it is important for archives to preserve data as completely and coherently as possible, maintaining critical information about context and connections. *NSIDC will use accepted standards in its data management practices. Data sets archived at and distributed by NSIDC will conform to NSIDC standard formats (see [Appendix C: Standards for Metadata, Documentation and Data Formats](#)). If data to be accepted by NSIDC are not in a preferred format, the relevant program or designated point of contact should review the costs of conversion to a preferred format to ensure consistency, availability and long-term access. The PCB and DD will regularly review these standards and update them as needed with advice from the Service Group Managers (see also section 1.4, Change Control).*

## 5.0 Data Set Documentation

"FD folders" historically have been used by NSIDC to capture relevant preservation information as well as

signed documentation and agreements about archived data sets where they exist. These folders *will be stored in and managed by Archival Services and appropriate information in them will be incorporated into the NSIDC metadata database. New data sets will continue to receive an FD folder accession number upon receipt at NSIDC and any relevant non-digital documents will be stored in them.*

*Where resources and agreed to level of service permits, NSIDC produces two types of data set documentation for published data sets, based on information that PIs submit with their data: (1) a Data Interchange Format (DIF) file and (2) a summary document.* Both standardize a data set's metadata, increase ease of use, and allow future access to that data set. The DIF is the metadata file format used by the NASA Global Change Master Directory (GCMD) and is compatible with International Standards Organization (ISO) and Federal Geographic Data Committee (FGDC) metadata standards. *NSIDC submits every DIF it writes to the GCMD, which is part of the U.S. Global Change Research Program.* The summary documents contain metadata for each data set, and are often much more comprehensive than the DIFs, providing detailed information about data collection methodology, file structure, tools for accessing data, and other data characteristics. They also describe how to appropriately cite the data set.

## 6.0 Archive Policies

### 6.1 Analog data:

Unlike digital data, value is ascribed to original analog data that is not given to copies. As such, *NSIDC recognizes that it is important for original analog data to be archived in the best conditions possible to preserve them for perpetuity, and that only copies be made available for users whenever possible.*

Users either travel to NSIDC to use these data or pay for the creation of a usage copy that can be shipped to them. To increase access to material, usage copies may be provided in formats other than the format of the original, as when a digital image is provided to users instead of a glass-plate negative. Currently, NSIDC has no budgeted resources to make usage copies of all analog data archived here. However, *analog data sets accepted from this point forward should include proper funding to address this need as well as proper storage, or should point to possible funding resources.* Additional funding is actively being pursued for all currently held data sets.

#### 6.1.1 Data Recovery efforts:

*In the event of a natural or man-made catastrophe that damages any analog data sets archived at NSIDC, the Information Center, working with Archival Services, will immediately implement data recovery efforts. These efforts will attempt to salvage and repair as much material as possible, especially of original archival material.*

### 6.2 Digital data:

#### 6.2.1 Data Capture Policies

*An offsite copy will be made for all digital data sets for which NSIDC is the primary archive and for other data sets as warranted by circumstances and as funding permits, as soon as practical upon ingest. Consistent with NSIDC standards, if such a data set is modified prior to publication, an offsite copy will be made of the modified data set and documentation describing the modifications made.*

## 6.2.2 Recovery Strategies

*To ensure data survival in the event of a catastrophe and consistent with industry best practices, NSIDC will maintain recovery strategies for data sets and associated preservation metadata where NSIDC is the primary archive, and for other data sets as warranted by circumstances and as funding permits.*

*Acceptable recovery strategies include:*

- *storing backup copies of the data in a secure, environmentally controlled offsite location as per federal guidelines;*
- *securing written agreements with external organizations to provide copies of the data on request, either by reprocessing data from a lower level product or by duplicating data from their archives.*

## 6.2.3 Media Management Policies

*NSIDC conducts ongoing archive media testing and refresh programs to assess the viability of its media archives and to refresh individual media on an as-needed basis. As media and archive technologies become obsolete, new technologies are chosen based on reliability and performance criteria, recognizing that cost may be an issue.*

# 7.0 Architecture and Security

Because of NSIDC's diverse agency funding, NSIDC currently does not have a center-wide technical infrastructure. *Common usage patterns have become established as defacto policy*, primarily based on DAAC infrastructure needs, as described below.

## 7.1 Current NSIDC Architecture

### Adcc:

The public FTP interface for ARCSS data housed at NSIDC. Unix-based

### Bipolar:

NSIDC's main web server used by NSIDC staff as well as the general public. Bipolar is the "production" or "live" server for the NSIDC internet, the NSIDC extranet, the DAAC Alliance site, and the ARCSS/ADCC virtual host <http://arcss.colorado.edu/>. Bipolar is also currently the development and production server for the NSIDC intranet (Igloo), though this site is in the process of being transitioned to Icemaker. Unix-based

### Icemaker:

NSIDC's development server and intranet web server. The intranet server is in the process of being transitioned from Bipolar. Location for Center-wide disk common space, shared to all other systems and users. Location for Center-wide common developer tools, shared to limited other systems. Icemaker is used by NSIDC staff who need to develop web sites and products. Unix-based

### Kryos:

E-mail, calendar, and printer server used by all NSIDC staff. Unix-based

### **N0\* machines:**

These machines, under the control of Raytheon via the ECS/EMD contract, hold the majority of NSIDC's archived data. Unix-based

### **Timberwolf:**

Robotic archive repository host that is used primarily by Operations and a few other NSIDC personnel. Unix-based

### **Sidads:**

The NSIDC public FTP interface. The system is used for data transmission between the Center and outside world. Access is limited by user accounts in most cases, though there is limited anonymous access for outgoing data/information. For use by anyone, though restrictions are noted when a user logs on. Unix-based

### **Snow:**

Lab-wide processing server. Used by NSIDC personnel for data processing. Unix-based

### **Glacier:**

Used to host tools that are used by the public and NSIDC staff. (e.g., GISMO). Presently is a disk server to other servers, but will be transitioned out of this function. Unix-based

## **7.2 Information Technology Security**

NSIDC's information technology related security policies are contained in the documents "NSIDC Information Technology Security Plan". Mike Stowe. December 2002 and "[NSIDC Information Technology Supplemental Security Policy](#)", Mike Stowe (undated)."

## **8.0 Data Tools**

*NSIDC supports and develops various tools to assist users in accessing, ordering, and using data according to agreed-upon levels of service and as resources permit. NSIDC publishes these tools on an [Internet web page](#). Requests for help with or questions about tools go through NSIDC User Services.*

## **9.0 Data Deletion and Retirement**

*The possibility of future data retirement or deletion must be covered with the contributing agency or individual at the time of data donation. Procedures for disposition of new data sets and reasons for retirement must be in writing and signed off by both the contributor and an NSIDC representative. Policies may differ according to program. Programs are encouraged to negotiate agreements with their sponsors and/or contributors to develop data deletion procedures for data held at NSIDC. Reasons for*

retirement of data might include data that have no historical value or data that have been archived for the long-term at another repository.

## 10.0 Data Access Policies

*NSIDC makes its data as widely available as possible and as quickly as possible. NSIDC follows best practices and adheres to program requirements. NSIDC data sets can be accessed through the online catalog or by contacting the User Services Office. NSIDC also contributes metadata to the GCMD and NOAA data directories, thereby increasing the avenues by which users can find data.*

### 10.1 Data charges

*Gaining access to or copies of some NSIDC products may require payment. Charges are made clear to potential users in data set documentation.*

### 10.2 Distribution Restrictions

#### 10.2.1 Data Provider

*Restrictions to data access will be negotiated between NSIDC and the data provider at the time the data are deposited at NSIDC. In general, time limits are placed on such restrictions to maintain NSIDC's policy of free access.*

#### 10.2.2 Non-Commercial Use Agreements

Some NSIDC data come with restrictions that limit their use to non-commercial purposes. *This information is provided in the data set documentation and is made clear to users before they are provided with data.* Generally a user must complete a form agreeing to these limitations before access is permitted.

#### 10.2.3 Copyright

Most of the data held by NSIDC are government-funded and in the public domain. *NSIDC archives some data for which it has no distribution rights. NSIDC makes copyright restrictions clearly known to potential users in the data set documentation, and refers the user to the copyright owner for additional rights.*

#### 10.2.4 Program Data

Depending on the program, NSIDC data may have access restrictions based on criteria such as time and/or membership (e.g., XYZ data is restricted to science team members only for the first year). *NSIDC works with the program to ensure that these access restriction requirements are met.*

## 11.0 References

NOAA and NASA provide the following documents on Data Management. Some guidelines from these documents were used in formulating this document. They are included here for your information.

[NASA Earth Science Enterprise Statement on Data Management](#)

[NOAA ESDIM Data Management document](#)

---

[1] Miller, Fredric M. 1990. Arranging and Describing Archives and Manuscripts. Chicago: Society of American Archivists, pp.3-4.

[2] O'Toole, James M. 1990. *Understanding Archives and Manuscripts*. Chicago: Society of American Archivists, pp.58-60.