



Levels of Service

Authors: R. Duerr, A. Leon, D. Miller, D.J Scott
Date 7/31/2009

CHANGE LOG

Revision	Date	Description	Author
1.0	7/31/2009	Original draft	Duerr, Leon, Miller, Scott
2.0	3/19/2010	Updated Documentation States	Leon, Miller, Duerr

Table of Contents

1. INTRODUCTION	4
1.1 PURPOSE AND SCOPE	4
2. SERVICE CATEGORIES	4
3. DEFINED LEVELS OF SERVICE	10
4. LEVELS OF SERVICE - STATE DESCRIPTIONS	5
4.1 ARCHIVAL STATES.....	5
4.2 METADATA STATES	5
4.3 DOCUMENTATION STATES	5
4.4 DISTRIBUTION STATES.....	5
4.5 USER SERVICES INFRASTRUCTURE STATES	7
4.6 USER SERVICES STATES.....	7

1. Introduction

1.1 Purpose and Scope

The purpose of this document is to:

- Define consistent terminology with which to discuss the variety of services that can be provided for a data set,
- Provide mechanisms for assessing and describing the relative level of effort to achieve a specified set of services for a data set, and to
- Assess the level of effort required to or released by moving a data set from one set of services to a different set of services for the data sets held and made available from the NSIDC data repository.

A team that included the leads of the three main groups that currently support dataset operations and maintenance, the technical writers group, the operations group, and user services developed this document. As a result, the document reflects both the organizational structure and technical systems in place at the time of writing.

In addition, it should be noted that a wide variety of services related to product generation also occur at NSIDC. However, these services are neither described nor discussed here. This document only describes services required for NSIDC to ingest, archive, and distribute data. Given the rapid pace of technological change, the expectation is that the terminology and levels of effort defined herein will change as new data services become available and NSIDC's system and organization evolves.

2. Service Categories

At this point in time six categories of services have been defined:

1. **Archival** – Levels of service in this area reflect the relative amounts of work required in order to ingest and archive a data set.
2. **Metadata** – Levels of service in this area reflect the amount of work required to develop and maintain metadata not just for the data set as a whole; but also for any data element or service associated with the data set.
3. **Documentation** – Levels of service in this area reflect the amount of work required to document the data set and any associated web pages.
4. **Distribution** – Levels of service in this area reflect the amount of work required to support data distribution or distribution-related services.
5. **USO Infrastructure** – In order to provide any level of User Services Office (USO) support, a base level of infrastructure support is required in order to prepare USO systems for the data set. This is a one-time setup that is not affected when levels of service change.
6. **USO Support** – Levels of service in this area reflect the amount of work required to support human-human requests for information about or help with a data set (via any mechanism – phone, email, etc.)

3. Levels of Service - State Descriptions

Within each category, a variety of different states or service levels can be provided. The states that are currently supported are defined below.

3.1 Archival States

PI site – Data is archived (and typically distributed from) the PI’s home site not NSIDC. Many of the PARCA data sets fall into this category along with all brokered products.

NSIDC 1-time ingest – The entire data set, be it a multiplicity of directories with multiple files or a single file, is delivered to NSIDC at one time. No updates are expected. The Greenland Drainage Basins data set is an example.

NSIDC periodic ingest – Ongoing data set updates are expected on an ad hoc, periodic basis – typically yearly. SSM/I data from RSI would be one example.

NSIDC ongoing ingest – Ongoing data set updates are expected on a routine automated scheduled basis – typically daily. All of the ECS data sets—AMSR-E, GLAS, MODIS, NISE—fall into this category.

3.2 Metadata States

None – Neither the provider nor NSIDC develop or maintain metadata about this data set. Many of the orphan data sets are in this state.

Skinny Catalog – Only the 8 mandatory fields from the NASA DIF are provided. This roughly corresponds to the Level 0 metadata category from the writer’s web pages.

Full Catalog – A complete catalog entry corresponding to the Level 1 metadata category on the writer’s web is provided for this data set. All of the products in the ECS system have full catalog metadata.

Inventory – Not only does the data set have a full catalog entry, but information is also kept about each and every file/granule within the data set. Data sets in the ECS system or Searchlight currently fall into this category.

Inventory Plus – Not only does the data set have a full catalog entry and a complete inventory, but additional metadata such as that necessary to provide or describe advanced data services is also available.

3.3 Documentation States

None – No documentation is provided. Many of NSIDC’s orphan data sets currently fall into this status.

Advertised Only – The data set is not in NSIDC’s catalog but is still advertised on an NSIDC Web page, such as a product Data Summaries Web page. No additional data set documentation beyond the Web page is available. Several PARCA data sets are in this state. This is a level of service we do not suggest using if at all possible. At minimum, we would like all of our data to be in our catalog.

Advertised + PI Provided Documentation – This is commonly used for data sets where usage is expected to be low or where resources for development of guide documentation are not available. A catalog page is created for this data set, and it is also advertised on an NSIDC Web page, such as a product Data Summaries Web page, but only the PI provided documentation is used. This documentation can vary depending on what the PI is providing.

1. If the PI provides a published paper that **describes the data well**, then the technical writer follows one of the three scenarios listed below:
 - If the PI has a published paper, and the paper has no copyright issues, a PDF of this paper is created and linked to from the catalog page. Example data set: *HDF4 Data Used to Assess Long-Term Access to Remote Sensing Data with Layout Maps, NSIDC-0397*.
 - If the PI has a published paper, but the paper has copyright issues and a Digital Object Identifier (DOI) Number, then the paper is linked to the DOI Number from the catalog page. Example data set: Hydrologic Sub-basins of Greenland, NSIDC-0371.
 - If the PI has a published paper, but the paper has copyright issues and no DOI number, we can’t legally distribute the paper. We can give a citation to the paper so others can find it on the Web, but we can’t supply it. Example data set: none currently.
2. If the PI’s published paper **doesn’t describe the data well**, or the PI **doesn’t have a published paper**, then the technical writer follows the scenario below:
 - We have the PI fill out Appendix C: Submission Template, if DAAC, or an equivalent if non-DAAC. We then create a PDF of this submission template and link to it from the catalog page. Example data set: Co-Registered AMSR-E, QuikSCAT, and WMO Data, NSIDC-0450.

Advertised + Mini-Guide – A catalog page is created for this data set, and it is also advertised on an NSIDC Web page, such as a product Data Summaries Web page, plus a streamlined NSIDC Guide document is developed for the data set and it is linked to from the catalog page. The following are the content criteria for a mini guide doc:

1. Data Set Title
2. Summary

3. Parameters
4. Data Set Citation
5. Investigator/Personnel
6. Temporal Coverage/Resolution
7. Spatial Coverage/Resolution
8. Format
9. References

Advertised + Full-Guide – A catalog page is created for this data set, and it is also advertised on an NSIDC Web page, such as a product Data Summaries Web page, and a full guide document is developed for the data set. Most ECS data products fall into this category. The criteria for a full guide document constitutes filling in all fields of the guide document template. However, there may be instances where some information is not available to completely fill in all the fields.

3.4 Distribution States

PI-site – The data are distributed from the PI's site. This typically is used for brokered data where NSIDC does not hold a copy of the data.

FTP access – Data are staged to a permanent FTP area for users to access. Many NSIDC DAAC non-ECS products fall into this category currently.

FTP + NSIDC granule level access – Data may still be available directly from an FTP area, but in addition granule level search and retrieval services are available via mechanisms such as WIST, the Data Pool search interface, Searchlight, etc.

FTP + NSIDC advanced services – Not only are the data available directly via FTP and through search and retrieval mechanisms, but a variety of advanced services (e.g., subsetting, reformatting, OGC services, etc.) are available.

3.5 User Services Infrastructure States

Infrastructure Preparation and Support – There is work required to prepare the USO infrastructure to support a data set, any data set, at any level of service. These activities include things like compiling and maintaining PI/referral contact information; creating canned responses (e.g., for order form registrations, etc.); updating the USO support tool with product information; and the like.

3.6 User Services States

Referral to PI or External Site – USO refers user questions to the PI or another external site.

USO documentation support – USO uses NSIDC documentation in an attempt to resolve user questions. USO often points user to area in documentation with information.

USO technical support – USO uses technical training on this data set to resolve user questions about ordering the data, extracting the data using available tools.

NSIDC technical support – USO uses all available resources (documentation, NSIDC programming staff, NSIDC scientists etc.) to resolve user questions. NOTE: This may even include consulting with external sources to resolve user questions.

4. Multiplicative Factors that can affect a State

Simply having the same level of service or state in a category does not guarantee that the amount of work required for each data set is the same. A number of factors may affect the work calculation. In general such factors are multiplicative. For example, they may make achieving the state “twice as much effort” or “half as much effort”. The factors that have been currently defined are listed in the table below. As with categories and states, the list of factors is expected to change over time.

Table 1 Multiplication Factors Affecting Work Difficulty

Category	Affecting Factors	Factor	States Affected
Archival	Data provided on standard media	1	All
	Data provided on obsolete media	5	All
	Use existing script for routine ingest	2	NSIDC ongoing ingest
	Develop script for routine ingest	3	NSIDC ongoing ingest
	Cursory 1-time QA	1	All
	Detailed 1-time QA	2	All
	Ongoing manual QA	4	NSIDC yearly ingest; NSIDC ongoing ingest
	Ongoing automated QA, including code development	3	NSIDC yearly ingest; NSIDC ongoing ingest
	No format/structure modification	0.75	All
	Minimal 1-time format/structure modification	1	All
	Significant 1-time format/structure modification	2	All
	Ongoing manual format/structure modification	4	NSIDC yearly ingest; NSIDC ongoing ingest
	Ongoing automated format/structure modification, including code development	3	NSIDC yearly ingest; NSIDC ongoing ingest
	Volume covered by current archive capacity	1	All
	Volume requires archive growth	3	All
	Requires new, independent archive	5	All
Standard NSIDC archive backup	1	All	

	copy		
	Additional and/or unique archive backup copy	3	All
Metadata	No available documentation	5	All, except none
	Difficult data formats	3	Full catalog, Cat + Inventory, Cat + Inventory Plus
Documentation	PI-provided documentation is complete and accurate	.5	All
	PI-provided documentation is incomplete but accurate	1	All
	PI-provided documentation is complete but inaccurate	3	All
	PI-provided documentation is incomplete and inaccurate	4	All
	No input from PI	5	All
	PI hard to reach	3	All
	Needs graphics, equations, or charts to explain the data	2	Full-Guide doc
	Data product needs own web site	4	All
	Complicated data set	2	All
Distribution	Permanent FTP distribution	1	All
	On-demand, temporary FTP distribution	2	All
	Distribution on media	4	All
	1-time pre-distribution manual processing	2	
	On-demand manual pre-distribution processing	4	All
	On-demand automated pre-distribution processing with existing code	2	All
	On-demand automated pre-distribution processing, including code development	3	All
User Services Support	Complex data set	2	All except Referral to PI or external site
	Multiple access paths	1.5	All except Referral to PI or external site
	NSIDC related tool support required	1.5	All except Referral to PI or external site
	Complex data format not supported by NSIDC	3	USO or NSIDC technical support
	PI-provided tools	2	USO or NSIDC technical support
	Missing data set information	2	All except Referral to PI or external site
	PI contact information is incorrect	1.5	Referral to PI

	Data set outside NSIDC expertise	3	All except Referral to PI or external site
	Monthly user support expected	.5	All
	Weekly user support expected	1.5	All
	Daily user support expected	3	All

5. Levels of Service Calculations

The following table (Table 1) provides a relative weighting of the amount of work that is normally required to take a data set to a particular state in each service category. To determine the total relative amount of work required for a particular data set, determine the score in each service category, determine whether there are any modifying factors (see Table 2) and apply them to the Work Difficulty Value for the appropriate service category, and sum the results for each category together. The larger the result the greater the relative cost to NSIDC. As a first cut, total scores between 0-60 can be considered to be low cost, while total costs greater than 100 should be considered high cost¹.

As the data management field in general and the data management knowledge of the user community increases over the upcoming years we can expect that the amount of information the user community is ready, willing, and able to provide will increase. In addition, changes in technology will undoubtedly induce NSIDC to add additional service levels. As a result, the table below should be considered a living document, one that should be reviewed and updated every few years at a minimum. Moreover, the simple 5, 10, 15 weighting scale was chosen to facilitate the addition of new levels between existing levels and should not be considered to be a linear estimate of effort required to move from level to level

Table 1: NSIDC Levels of Service Categories and States

			GENERIC EXAMPLES			
Category	Desired State	Work Difficulty (see footnote)	Brokered	Distribution in Bulk	Distribution by Granule	Distribution by Granule with Services
Archival	PI site	0	Y	N	N	N
	NSIDC 1-time ingest	5	N	M	M	M
	NSIDC yearly ingest	10	N	M	M	M
	NSIDC ongoing ingest	15	N	M	M	M

¹ The Work Difficulty scale is a relative and subjective measure of the amount of

Metadata	None	0	N	N	N	N
	Skinny Catalog	5	Y	M	N	N
	Full Catalog	10	M	Y	N	N
	Cat + Inventory	15	N	N	Y	N
	Cat + Inventory Plus	20	N	N	M	Y
Documentation	None	0	M	N	N	N
	Advertised Only	5	Y	N	N	N
	Adv + PI-Provided	10	M	Y	M	M
	Adv + Mini-Guide	15	M	M	Y	M
	Adv + Full-Guide	20	M	M	M	Y
Distribution	PI site	0	Y	N	N	N
	FTP Access	5	N	Y	N	N
	FTP + NSIDC granule level access	10	N	N	Y	N
	FTP + NSIDC advanced services	15	N	N	M	Y
USO Infrastructure	Infrastructure Preparation	5	Y	Y	Y	Y
User Services Support	Referral to PI or external site	5	Y	M	M	M
	USO documentation support	10	M	Y	M	M
	USO technical support	15	M	M	Y	M
	NSIDC technical support	20	N	N	M	Y
Score Range			15 to 45	30 to 65	45 to 85	55 to 90
Example			Parca Data			ECS Data (MODIS, AMSR-E, GLAS)

In the above table, Y=yes, N=no, M=maybe, NA or a blank cell=not applicable.

5.1 Calculating Relative Effort to Archive a New Data Set

5.1.1 Example 1: Simple FTP Distribution Support:

A PI requests NSIDC to archive and distribute the data she acquired during a particular field campaign. The data is in a standard format and is available electronically (FTP pull) but is expected to be of interest to a limited set of users, so only a skinny catalog entry and readme file are scoped for development. Fortunately, the PI is willing to provide the metadata and information for the readme file. NSIDC expects to distribute the data via an FTP site and provide basic documentation support for end users.

Archival State = NSIDC 1-time ingest =	5	
Affecting factors: No format mods are needed =	x.75	
		Archive total: 3.75

Metadata State = Skinny catalog =	5	
Affecting factors: none	x 1	
		Metadata total: 5.0

Documentation State = Adv+Read me =	10	
Affecting factors: PI willing to provide	x.5	
		Documentation total: 5.0

Distribution State = FTP Access =	5	
Affecting factors: none	x1	
		Distribution total: 5.0

User Services Infrastructure Preparation		5.0
--	--	------------

User Services State = Documentation support	10	
Affecting factors: Monthly user support expected	x.5	
		USO Services total: 5.0

Grand Total 28.75

5.1.2 Example 2: Complex time series with web service support

A PI requests NSIDC to ingest and distribute a 10-year time series of daily observations. The observations are ongoing and the PI would like NSIDC to acquire them as soon as they are available after being QC'd by the PI. The PI had developed a set of OGC services with which to access the data and would like NSIDC to continue supporting these services. Given the sizeable user community the PI has been supporting NSIDC agrees to do so. However, the data set is complex, not all of which NSIDC has experience with, and the PI does not like writing documentation.

Archival State = NSIDC ongoing ingest =	15	
Affecting factors: Need to develop ingest script	x3	
		Archive total: 45
Metadata State = Catalog + Inventory Plus	20	
Affecting factors: None	x 1	
		Metadata total: 20
Documentation State = Adv+Full-guide =	20	
Affecting factors: Limited PI input	x 5	
Complicated data set	x 2	
		Documentation total: 200
Distribution State = FTP + NSIDC advanced services =	15	
Affecting factors: Automated pre-distribution code needed	x3	
		Distribution total: 45
User Services Infrastructure Preparation		5.0
User Services State = NSIDC Technical support	20	
Affecting factors: Daily user support expected	x3	
Complex data set	x2	
Multiple access paths	x1.5	
		USO Services total: 180
		Grand Total 495

5.2 Calculating the Effort to Change the Level of Service

5.2.1 Raising the Level of Service for a Data Set

Due to some change in climate conditions a data set that is updated once a year, which had been primarily of interest to researchers in a specific sub-discipline suddenly becomes popular to a broad user community (as evidenced by a huge increase in user requests and associated requests for support). Given this increased interest the PI wishes to hand off the increased user support to NSIDC. Consequently NSIDC needs to consider raising the level of support. In addition, NSIDC decides to assess 1) decreasing the lag time before new data is available to users, 2) improving the metadata, documentation and other information that is available about the data, and 3) providing the data in the form and through services that the new user community seems to be demanding. The PI agrees to provide NSIDC the data as soon as it is available (rather than yearly) in exchange for USO user support.

Initial Levels of Service:

Archive: NSIDC yearly ingest	10
Metadata: Skinny Catalog	5
Documentation: Advertised + Readme	10
Distribution: FTP Access	5
USO Infrastructure:	5
User Services Support: Referral to PI	5
Total	40

Desired Final Levels of Service:

Archive: NSIDC ongoing ingest ²	15
Metadata: Skinny Catalog	20
Documentation: Advertised + Readme	20
Distribution: FTP Access	15
USO Infrastructure:	5
User Services Support: Referral to PI	20
Total	95

Relative increase in Levels of Service, i.e., 95 – 40 = 55

In other words, assuming there are no additional complicating factors, the amount of work that NSIDC needs to do to raise this data set to the new level of service is greater than the amount of work NSIDC had to do to originally achieve the existing level of service.

² Assuming that the PI agrees to provide NSIDC the data as soon as it is available (rather than yearly) in exchange for USO user support