

## Member Node Description: NSF Arctic Data Center

---

Version 1.0      3/23/2016      Matthew B. Jones

### General

**Name of resource:** National Science Foundation Arctic Data Center  
**URL(s):** <https://arcticdata.io>  
**Institutional affiliation(s):** NCEAS, UC Santa Barbara; NOAA NCEI; DataONE  
**Primary geographic location:** Santa Barbara, CA, USA  
**Project Director & contact info:** Matthew B. Jones [jones@nceas.ucsb.edu](mailto:jones@nceas.ucsb.edu)  
**Technical Contact & contact info:** Matthew B. Jones [jones@nceas.ucsb.edu](mailto:jones@nceas.ucsb.edu)  
**Age of resource:** Since 2016, but with prior history back to 2008  
**Funding support:** NSF  
**Proposed Unique Identifier:** urn:node:ARCTICA

### Content

**Content description/collection policy (1 paragraph, domain and spatial/temporal coverage, uniqueness of content, exclusions, as applicable):**

Data and software products of NSF-funded Arctic program research grants. Data coverage spans international boundaries across the Arctic region, and currently includes data covering time periods from 1800 to present. Disciplines span the earth and environmental sciences, and include social science, policy, and economics research from the Arctic region.

**Types of data (complex objects, text, image, video, audio, other):**

Tabular, spatial raster, spatial vector, image, video, and audio data are included in the collection.

**Data and metadata availability (rights, licensing, restrictions):**

Data and metadata generally available under CC-0 or CC-BY license, except in cases of sensitive data which may be restricted access under special exceptions, especially for human subjects data.

**Option for embargo (yes/no, duration):**

Yes, upon request, but typically data are released upon submission. NSF Arctic programs require rapid data release for all funded projects, so long embargos would be exceptional.

**Size of holdings (number and size of datasets, mean and median granules (files) per dataset):**

Currently, over 3500 data sets comprising 486,000 objects are in the collection, with an average of 149 files per package, and a median of 4 files per package. Total size of the holdings as of March 2016 is 4.7TB.

**Please describe recent usage statistics, if known, including information on annual data product downloads, annual number of users, annual number of data products used in publications:**

Approximately 450 users contributed data to date. Data have been used in a wide variety of publications, which are compiled online at: <http://arcticdata.io/publications>

## User interactions

### **How does a user contribute data? (what can be deposited, how are data prepared, are specific software required, documentation/support available)**

Data can be contributed by any researchers funded by NSF Arctic programs, either currently or historically. Data are prepared by contributing scientists, and can be uploaded via a web-based entry system, through tools such as R and Matlab, or via a REST API. Data submission documentation is available online at <https://arcticdata.io/submit>.

### **How does a user acquire / access data?**

Data access is through the search and browse portal at <https://arcticdata.io>. Search filters based on contributor, subject, and spatial, temporal, and taxonomic coverage are provided.

### **What user support services are available (both for depositing and accessing/using data)?**

User support is available to NSF researchers for data access, data submission, and data management planning. The Arctic Data Center maintains a support email channel at [support@arcticdata.io](mailto:support@arcticdata.io), as well as live chat systems on the web site.

### **How does the resource curate data at the time of deposit?**

Metadata and data are reviewed by the Center's curatorial team, and requests for corrections are made to the contributing researchers. When the data are ready for deposition, Center staff mint a DOI and make the data set publicly accessible. Plans are in place to add automated metadata and data quality checking systems to help researchers prepare and maintain high-quality products.

## Technical characteristics and policies

### **Software platform description, incl. data search and access API(s):**

Metacat, MetacatUI for search, and the DataONE REST API

### **Service reliability (including recent uptime statistics, frequency of hardware refresh, if known):**

No statistics available, service is commencing in 2016.

### **Preservation reliability (including replication/backup, integrity checks, format migration, disaster planning):**

Data will be preserved via replication to multiple off-site locations within the DataONE network, including our partner NOAA's NCEI data systems. In addition, a full backup of all data and metadata are maintained in the Amazon public cloud, as well as locally at UC Santa Barbara on both spinning disk and tape. Fixity of the collection is handled via periodic replication audits.

### **User authentication technology (incl. level of create/modify/delete access by users):**

Users authenticate via the ORCID (<http://orcid.org>) researcher identity system. Authorized users are granted write access to the system for contributing Arctic data.

### **Data identifier system and data citation policy, if available:**

Data identifiers are minted using multiple syntaxes. Top-level data sets are assigned DOIs, while individual files and granules within data sets are typically assigned UUIDs.

### **Metadata standards (including provenance):**

Many metadata standards will be accepted, but users are guided towards Ecological Metadata Language and ISO 19139 for scientific metadata. Some tools will use ProvONE for provenance relationships.

## Capacity/services to DataONE

At what functional tier will you initially be operating? (see <http://bit.ly/MNFactSheet> for definitions)

- Tier 1: Read only, public content
- Tier 2: Read only with access control
- Tier 3: Read/write using client tools
- Tier 4: Able to operate as a replication target

If you can host data from other member nodes, what storage capacity is available?

None without explicit agreements.

Can you provide computing capacity to the broader network? If so, please describe.

Computing capabilities are available to NSF Arctic researchers.

## Other Services

What other services or resources (such as expertise, software development capacity, educational/training resources, or software tools) can be provided of benefit to the broader network?

The NSF Arctic Data Center supports a variety of outreach and training activities for the Arctic research community, including bi-annual training events, outreach at Arctic research conferences, and the development of data management planning materials. Consulting information systems and infrastructure for specific Arctic-related research projects is possible under specific arrangements.