Member Node Description: International Arctic Research Network

General

Name of resource:	International Arctic Research Network (IARC) Data Archive
URL(s):	http://climate.iarc.uaf.edu/geonetwork
Institutional affiliation(s):	University of Alaska Fairbanks
Primary geographic location:	Fairbanks, Alaska, United States
Project Director & contact info:	Larry Hinzman Director, International Arctic Research Center lhunzman@iarc.uaf.edu
Technical Contact & contact info:	James Long jlong@iarc.uaf.edu
Age of resource:	Since 2010
Funding support:	Developed under DOE award number: DE-SC0005868
Proposed Unique Identifier:	urn:node:IARC

Content

Content description/collection policy (1 paragraph, domain and spatial/temporal coverage, uniqueness of content, exclusions, as applicable):
The IARC Data Archive is home to Arctic field and modeling data.

Types of data (complex objects, text, image, video, audio, other):
No restriction on data types.

Data and metadata availability (rights, licensing, restrictions):
Data and metadata are either freely available, or under embargo pending publication.

Option for embargo (yes/no, duration):
Yes, determined on an individual basis, usually not more than a few years.

Size of holdings (number and size of datasets, mean and median granules (files) per dataset):
About 5 TB of holdings.

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:
In the last 6.5 months, there were 66875 web page hits from 9657 unique IP addresses, representing 6365 class B networks.

User interactions

How does a user contribute data? (what can be deposited, how are data prepared, are specific software required, documentation/support available)
A web front end accepts user contributions at http://climate.iarc.uaf.edu/geonetwork, for which an account is required.

How does a user acquire / access data?
Data access and acquisition occur at http://climate.iarc.uaf.edu/geonetwork.

What user support services are available (both for depositing and accessing/using data)?
A web support page lives at http://data.iarc.uaf.edu/archive.html, and email support is available at archive@iarc.uaf.edu

How does the resource curate data at the time of deposit?
Data deposition is reviewed on a case-by-case basis.

Technical characteristics and policies

Software platform description, incl. data search and access API(s):
User interfaced is provided by GeoNetwork. http://geonetwork-opensource.org/ lists support for: "Metadata standards (ISO19115/ISO19119/ISO19110 following ISO19139, FGDC and Dublin Core), Catalog interfaces (OGC-CSW2.0.2 ISO profile client and server, OAI-PMH client and server, GeoRSS server, GEO OpenSearch server, WebDAV harvesting, GeoNetwork to GeoNetwork harvesting support) and Map Services interfaces (OGC-WMS, WFS, WCS, KML and others) through the embedded GeoServer map server."

Service reliability (including recent uptime statistics, frequency of hardware refresh, if known):
99+% uptime.

Preservation reliability (including replication/backup, integrity checks, format migration, disaster planning):
At least 3 copies of data exist within 24 hours of ingestion, and 5 copies within one week. One copy is kept off-site.

User authentication technology (incl. level of create/modify/delete access by users):
Authentication is via https login with username and password. Users may modify their data and metadata, deleted records are retained.

Data identifier system and data citation policy, if available:
md5 hash is used on the GeoNetwork side, sha1 on the DataONE Generic Member Node (GMN) side.

Metadata standards (including provenance):
iso19139 (xml version of iso19115) on the GeoNetwork side, dublin core extended on the DataONE Generic Member Node (GMN) side.

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?
We plan to eventually become a tier 4 member node, with initially 30 TB.

Can you provide computing capacity to the broader network? If so, please describe.
This is possible, via spare cluster cycles at lowest priority.

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?
We are interested in building semantic tools on top of the DataONE framework.

DataONE software written:

geo2d1.py - Tier 1 code to export GeoNetwork metadata into a DataONE GMN.
The International Arctic Research Center (IARC) Data Archive uses the xml
form of ISO 19115 metadata, ISO 19139. Not all researchers fill out
metadata that validates, however, so this code xslt transforms the public
ISO 19139 metadata records from IARC's geonetwork OAI-PMH endpoint into
dublin core extended (dcx), and uploads the validated dcx into IARC's
DataONE GMN as metadata, while the original ISO 19139 metadata is uploaded
as data (text/xml). A resource map relating the two completes the package.
Each time this script runs, it checks to see if a package update is required:
the current GMN ISO 19139 data object (xml) is compared with the downloaded
OAI-PMH version, and if different, triggers the package update.