

# INTERNATIONAL TRENDS IN POLAR DATA MANAGEMENT

Perspectives from the Standing Committee  
on Antarctic Data Management

Johnathan Kool –SCADM chair, Australian Antarctic Data Centre



# ABOUT SCADM

- Standing Committee on Antarctic Data Management
  - <https://www.scar.org/resources/scadm/overview/>
- Falls under the auspices of SCAR (<https://www.scar.org/>)
- Chair: Johnathan Kool (Australian Antarctic Division)
- Co-Vice Chairs:
  - Helen Peat (British Antarctic Survey)
  - Frank Nitsche (Lamont-Dohert Earth Observatory, USAP)
- Members from 33 countries (and counting)
- Monthly meetings



# OBLIGATIONS ON THOSE COLLECTING DATA IN POLAR REGIONS

Requirements under the Antarctic Treaty - Section III c

- *“scientific observations and results from Antarctica shall be exchanged and made freely available”*

Objective of the International Arctic Science Committee (IASC)

- *“seek to ensure that scientific data and information from the Arctic are safeguarded, freely exchangeable and accessible”*



Photo credit: NASA

# DATA POLICIES

- SCAR Data Policy - Currently undergoing revision

1

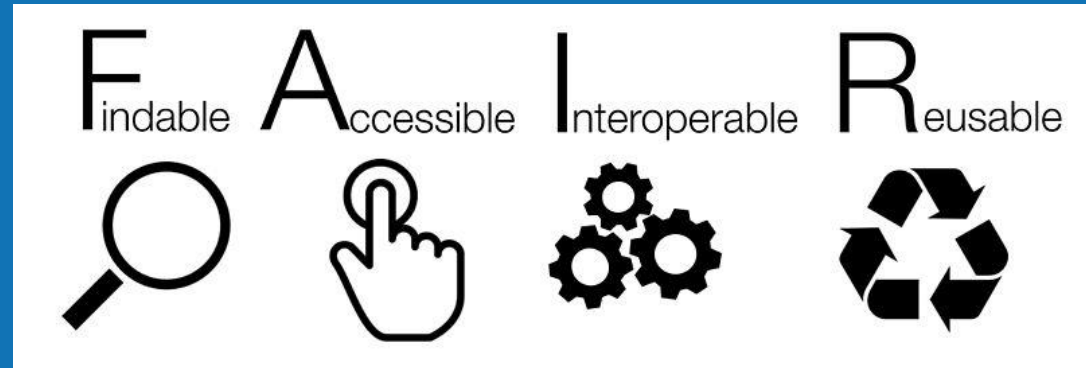
## Alignment of Polar Data Policies - Recommended Principles

**Draft**

*Contributors: Stein Tronstad (1), Pip Bricher (2), Johnathan Kool (3), Peter Pulsifer (4), Anton van de Putte (5), Jan Rene Larsen (6), Helen Peat (7), David Rayner (8), Taco de Bruin (9), Marten Tacoma (9), Jen Thomas (10), Frank Nitsche (11), VS Samy (12), Dariusz Ignatiuk (13), Erik Buch (14), Anne Treasure (15)*

*(1) Norwegian Polar Institute, (2) Southern Ocean Observing System, (3) Australian Antarctic Data Centre, (4) Carleton University, (5) Royal Belgian Institute of Natural Sciences, (6) Arctic Monitoring and Assessment Programme, (7) British Antarctic Survey, (8) Swedish National Data Service, University of Gothenburg, (9) Royal Netherlands Institute for Sea Research, (10) Swiss Polar Institute, (11) Lamont-Doherty Earth Observatory of Columbia University, (12) National Centre for Polar and Ocean Research, Goa, India, (13) Svalbard Integrated Arctic Earth Observing System, (14) European Global Ocean Observing System, (15) South African Environmental Observation Network (SAEON) Information Systems*

# MAKING DATA FAIR

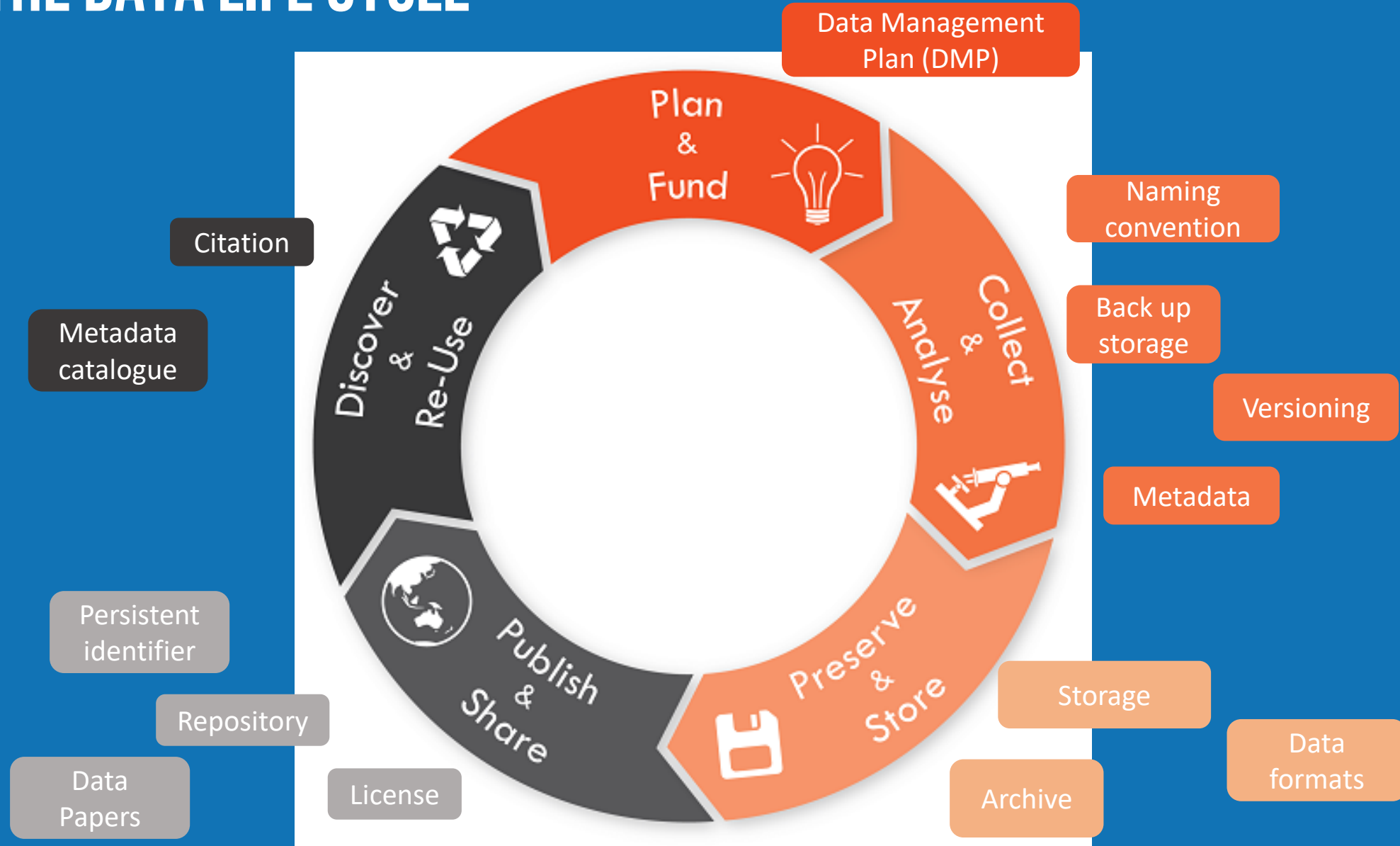


Perce 11

- FAIR – Findable, Accessible, Interoperable, and Reusable
- Findable: Easily discoverable – e.g. via Google, catalogues, linked websites
- Accessible: getting to the data – open downloads, data streams, APIs
- Interoperable: combinable data – standards, consistent formats, vocabularies
- Reusable: derived products – versatile, non-restrictive, clean, understandable



# THE DATA LIFE CYCLE



# DATA MANAGEMENT PLANS

- Assist with the process of meeting data management requirements as work progresses.
- What will you do with data during and after research
- Who, what, when, where, how, and why
- <https://www.ands.org.au/working-with-data/data-management/data-management-plans>

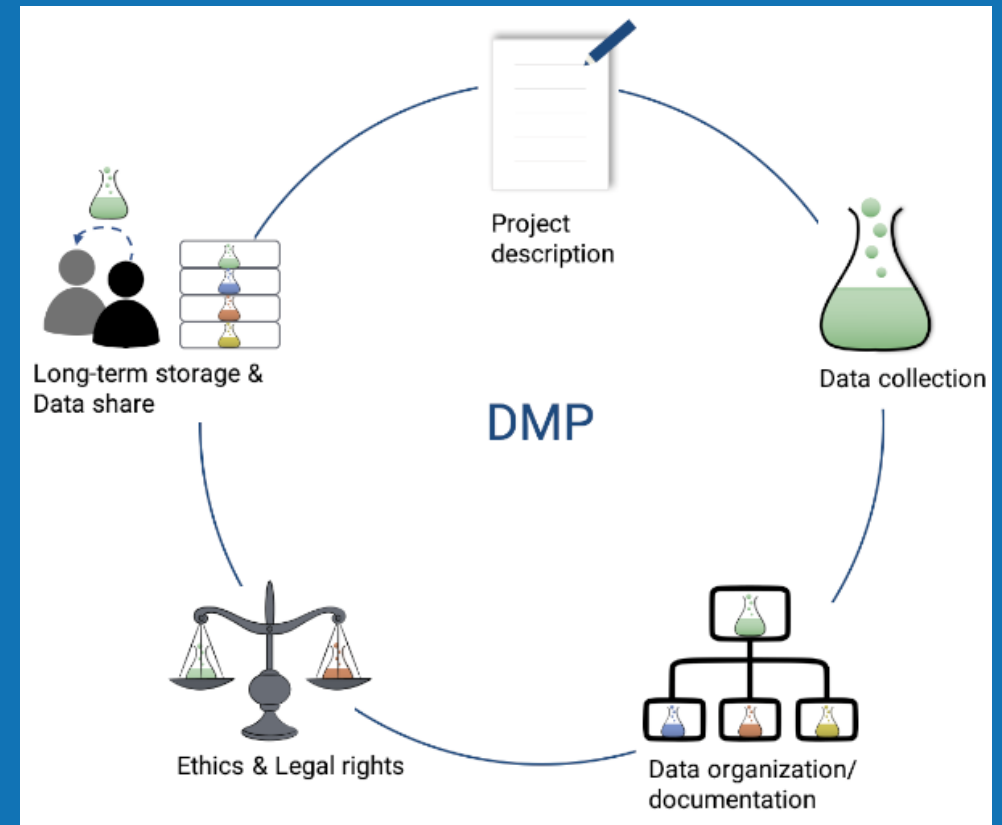
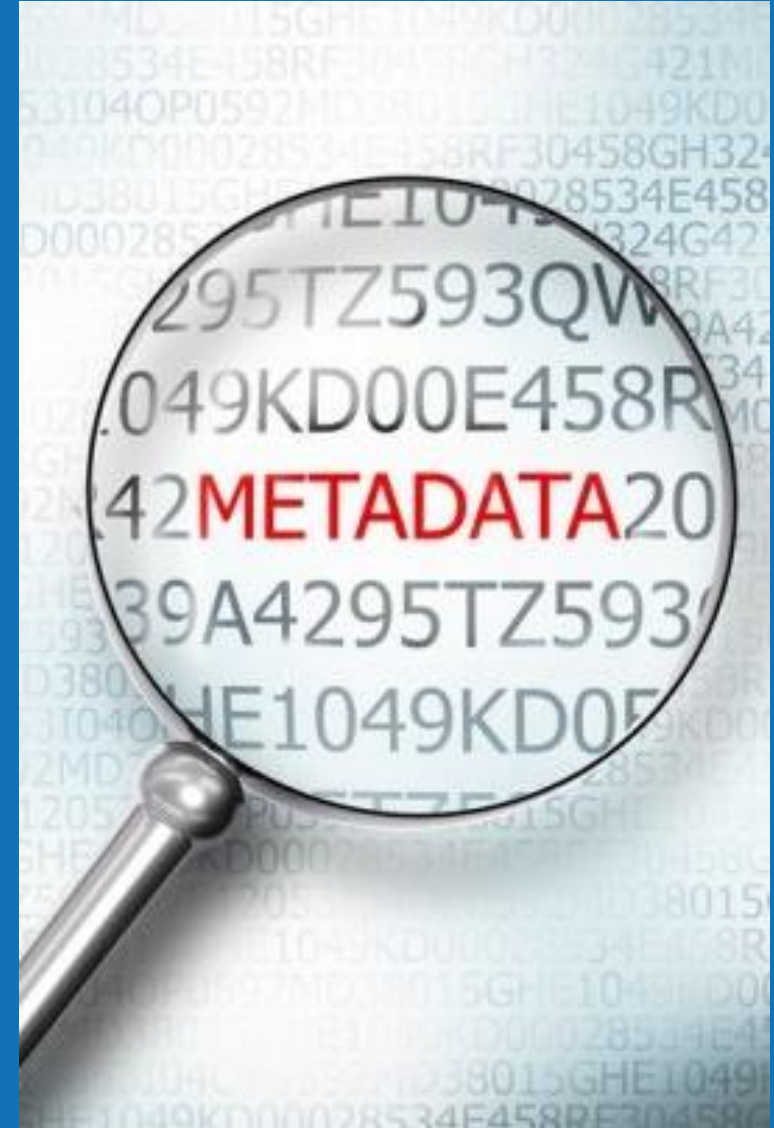


Image credit: Franziska Kappert

# METADATA

- Data about data – e.g.
  - Creator(s)
  - Keywords
  - Download information
  - Usage restrictions
- Basis of searchability (and also interoperability)
- DIF (Directory Interchange Format), ISO-19115, DCAT





# PERMISSIONS AND SHARING

- Data should be accompanied by some form of rights statement, public domain statement, or an internationally recognised data license.
- The absence of any of the above may make the data legally unusable in some jurisdictions!
- Rights statements – [rightsstatements.org](https://rightsstatements.org)
- Public domain statement (e.g. CC0)
- International data licenses – Creative Commons
- Open-source software – MIT, Apache, BSD-3



# ATTRIBUTION AND PROVENANCE

- Licensing – e.g. CC-BY requires attribution
- Persistent identifiers
  - DOIs for digital objects (data, databases)
  - ORCID for researchers (people)
- Snapshots in time and amalgamations can become complicated
  - Open StreetMap, Google Earth
  - Dynamic DOIs



# FINDING ANTARCTIC DATA

- Google
- Google Datasets
- SOOS – <https://www.soos.aq>
- Antarctic Master Directory/EarthData
  - <https://search.earthdata.nasa.gov/portal/amd/search>
- Your friendly neighbourhood Polar Data Centre – for example
- Australia – <https://data.aad.gov.au>
- UK – <https://bas.ac.uk/data>
- US - <https://www.usap-dc.org/>
- Pangaea - <https://www.pangaea.de/>

The screenshot displays the Earthdata Search portal interface. At the top, the browser address bar shows the URL [search.earthdata.nasa.gov/portal/amd/search](https://search.earthdata.nasa.gov/portal/amd/search). The page header includes the NASA logo, the text "EARTHDATA", and a search bar. Below the header, the search results are displayed, showing "8,835 Matching Collections". A sidebar on the left provides a "Filter Collections" menu with various categories and features. The main content area lists several data collections, including MODIS/Terra Near Real Time (NRT) Surface Reflectance and Thermal Anomalies/Fire data, and GPM MHS on NOAA-19 and METOP-A Radiometer Precipitation Profiling data. Each collection entry includes a thumbnail, a title, a description, and a "View" button. The footer of the page contains the version number "v1.153.6", search time "2.0s", and links to NASA Official, FOIA, and Privacy Policy pages.

Earthdata Search :: Antarctic Mas... X +  
search.earthdata.nasa.gov/portal/amd/search

EARTHDATA Find a DAAC -  
AMD SEARCH

Search for collections or topics

8,835 Matching Collections  
Showing 20 of 8,835 matching collections

Filter Collections

Categories

Features

- Available from AWS Cloud
- Customizable
- Map Imagery
- Near Real Time

Keywords

Platforms

Instruments

Organizations

Projects

Processing Levels

Data Format

Tiling System

Horizontal Data Resolution

MODIS/Terra Near Real Time (NRT) Surface Reflectance Daily L2G Global 250m SIN Grid  
Int'l / Interagency • 2015-12-06 ongoing • The MODIS Near Real Time (NRT) Surface Reflectance products are an estimate of the surface spectral reflectance as it would be measured at ground level in the absence of atmospheric scattering or absorption. Low-level data are corrected ...  
CWC NRT MOD09GQ v6NRT - NASA/GSFC/EOS/ESDIS/LANCEMODIS

MODIS/Terra Near Real Time (NRT) Thermal Anomalies/Fire 5-Min L2 Swath 1km  
2,040 Granules • 2015-12-07 ongoing • MODIS Near Real Time (NRT) Thermal Anomalies/Fire products are primarily derived from MODIS 4- and 11-micrometer radiances. The fire detection strategy is based on absolute detection of a fire (when the fire strength is sufficient to detect),...  
NRT MOD14 v6NRT - NASA/GSFC/EOS/ESDIS/LANCEMODIS

GPM MHS on NOAA-19 (GPROF) Radiometer Precipitation Profiling L2A 1.5 hours 17 km V05 (GPM\_2AGPROFNOAA19MHS) at GES DISC  
37,306 Granules • 2014-01-31 ongoing • Version 5 is the current version of the data set. Version 4 is no longer available and has been superseded by Version 5. The 2AGPROF (also known as, GPM GPROF (Level 2)) algorithm retrieves consistent precipitation and related science fields f...  
GPM\_2AGPROFNOAA19MHS v05 - NASA/GSFC/SED/ESD/GCDC/GESDISC

GPM MHS on METOP-A (GPROF) Radiometer Precipitation Profiling L2A 1.5 hours 17 km V05 (GPM\_2AGPROFMETOPAMHS) at GES DISC  
28,614 Granules • 2014-01-31 ongoing • Version 5 is the current version of the data set. Version 4 is no longer available and has been superseded by Version 5. The 2AGPROF (Goddard Profiling) algorithm retrieves consistent precipitation and related science fields from the following ...  
GPM\_2AGPROFMETOPAMHS v05 - NASA/GSFC/SED/ESD/GCDC/GESDISC

Looking for more collections? [Leave AMD's Search Portal](#)

v1.153.6 • Search Time: 2.0s • NASA Official: Stephen Berrick • FOIA • NASA Privacy Policy • USA.gov

**AUSTRALIAN  
ANTARCTIC  
DIVISION**

# RSV NUYINA CAPABILITIES

- **GPS and Positioning (MRU)**
- **Seabed Mapping**
  - Kongsberg EM 122 (Deep)
  - Kongsberg EM 712 (Shallow)
  - Simrad EK80 broadband echosounder
- **Fish Acoustics**
  - SimRad SH90 (downward) and MS70 (sidescan)
- **Radar**
  - Arc-X Doppler weather, Sigma 6 Ice Wave
- **Oceanography**
  - ADCP, OceanPack (7 variables), PCO<sub>2</sub>, ISAR SST
- **Moving Vessel Profiler**
  - Continuous CTD and SVP
- **Atmospheric**
  - Radiation sensors, Wind, Temperature, Pressure, Clouds
- **USBL Capability**
- **Flowthrough system**





Satellites



Data Sources



Field Scientists



Drones and Autonomous Vehicles



RSV Nuyina

Files



Data Storage

Cloud Services



Databases

Metadata

@Title  
@Author  
@Keywords  
@Fields

@Structure  
@Meaning

...

Web Services

Dataset search

Automated weather station data

Place names gazetteer

Data Interrogation & Processing

... and more

Analytics

$$y=f(x_1,x_2,x_3\dots)$$



AI



Reports



Data Products & Visualisation



Maps



Data Science Notebooks



Dashboards



# CLOUD SERVICES

What's up with 'The Cloud'?

- Scalability
- Accessibility
- Portability
- Cost effective (if done well)

Simple Storage Service (S3)

- Object store – similar to a file system, but not quite...
- Everything is accessible through code via a URL



# Remote Sensing of Near-Coastal Antarctic Sea Ice and Its Impacts on Ice Shelves and Ecosystems.

Massom, R.

ASAC\_3024

Massom, R. (2009) Remote Sensing of Near-Coastal Antarctic Sea Ice and Its Impacts on Ice Shelves and Ecosystems., Ver. 1, *Australian Antarctic Data Centre* - [https://data.aad.gov.au/metadata/record/s/ASAC\\_3024](https://data.aad.gov.au/metadata/record/s/ASAC_3024), Accessed: 2021-10-08



## Dataset: fast\_ice\_1997\_1999

| public | 30 Jan 2012

[Browse dataset](#)

[Request dataset](#)

## Quality

The values provided in temporal and spatial coverage are approximate only. Taken from the 2008-2009 Progress Report: Variations to work plan or objectives: The only variation has been to geographically extend Objective 3 to analyse the potentially significant impact of very thick fast ice on the dynamic and calving behaviour of the floating Mertz Glacier Tongue (see Section 1.1 above). It is anticipated that this extension will enhance the project, by highlighting a previously overlooked yet potentially important process affecting floating glacier tongue/ice shelf breakup, whereby strong coupling between very thick perennial sea ice and continental ice provides stability in certain ice sheet margin regions. It also enhances the link with AAS Project #2698 (PI: Warner), on which Dr Massom is a Co-I ("Antarctica - past, present, and future: exploring the dynamic interactions of ice sheet and ice shelves within the global climate system through computer modeling"). Field work: This is predominantly a remote sensing-based project. However, in situ and airborne observations of fast ice acquired during SIPEX (V1 2007/8) by other projects on which the PI (Massom) is a Co-I (#2901 and 3030) may prove to be useful in helping to interpret satellite laser altimeter-based estimates of fast ice thickness. Moreover, and given its extreme

## About this data record

Metadata record for data from AAS (ASAC) Project 3024. Public The proposed research will derive improved estimates of East Antarctic fast-ice extent and thickness, and their variability, from satellite data. These will be used to explicitly test relationships between fast ice/other environmental parameters and Emperor penguin population dynamics. We shall also combine observations with a wave-ice shelf-sea ice interaction model to test the hypothesis that catastrophic ice shelf break-up events on the E. Antarctic Peninsula are linked to...

[Show more](#) ▼

**Coverage:** 30 Sep 2008 - 31 Mar 2011

**Created:** 23 Apr 2009

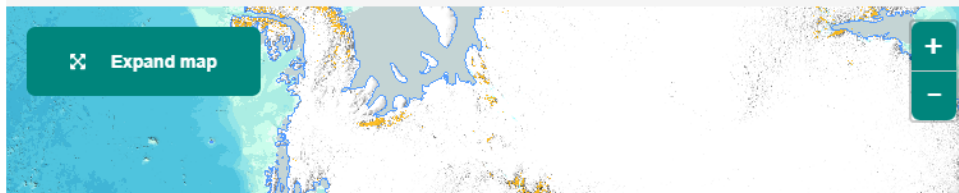
**Revised:** 26 Apr 2017

**Expected release date:** 8 Oct 2021

**Version:** 1

**Data centre:** [Australian Antarctic Data Centre](#)

Lat N -70.0 S -63.0 Lon W -180.0 E 180.0



## Fast-ice Distribution in East Antarctica During 1997 and 1999 Determined Using RADARSAT Data

An image correlation technique has been applied to RADARSAT ScanSAR images from November in 1997, and November 1999, to create the first detailed maps of fast ice around East Antarctica (75E-170E). This method is based upon searching for, and distinguishing, correlated regions of the ice-covered ocean which remain stationary, in contrast to adjacent moving pack ice. Within the overlapping longitudinal range of ~86E-150.6E, the total fast-ice area is 141,450 km<sup>2</sup> in 1997 and 152,216 km<sup>2</sup> in 1999. Calibrated radar backscatter data are also used to determine the distribution of two fast-ice classes based on their surface roughness characteristics. The outer boundaries of the determined fast-ice area for November in 1997 and 1999 are contained in the data files for this record. This work has been allocated to ASAC project 3024.

Search

Resource	Type	Last Modified	File Size
<a href="#">ASAC_3024.xml</a>		2020-05-22 00:46	
<a href="#">LICENSE</a>		2020-05-22 00:21	
<a href="#">README</a>		2020-07-02 12:06	
<a href="#">fast-ice/</a>	Folder		

4 records

[Download Dataset](#)

[View Full Record](#)

### Citation

Massom, R. (2012) Fast-ice Distribution in East Antarctica During 1997 and 1999 Determined Using RADARSAT Data, Ver. 1, *Australian Antarctic Data Centre* - doi:10.4225/15/5afe50d677c8e, Accessed: 2021-10-08

# THE BEGINNINGS OF BIG DATA

- All of the AADC's collections are available online
  - Online does not necessarily mean public
  - Datasets are organised at the top level according to level of protection (public, embargoed, AADC-only, etc.), but access is also controlled using permissions/policies.
- File size is no longer limiting
- Analytics can be performed “in-place”







# IMPROVED RELIABILITY

- Cloud services are also improving data resilience
- We have been working with IT to trial corporate backups of data to AWS.
- Our data collections are >100 TB - too large for our tape systems to handle effectively.
- AWS' systems have 99.999999999% reliability





# BUSINESS INTELLIGENCE AND GOVERNANCE

- Reporting, metrics
- Matomo, Google Analytics
- Logging
- ORCID

The screenshot shows the ORCID iD profile for Aleks Terauds. The profile includes the following information:

- ORCID ID:** <https://orcid.org/0000-0001-7804-6648>
- Employment (1):** Australian Antarctic Division: Kingston, TAS, AU (2011-04 to present | Principal Research Scientist (Antarctic Conservation and Management) Employment). Source: Aleks Terauds (Preferred source).
- Works (50 of 55):** Items per page: 50. 1 - 50 of 55.
  - Rapid condition monitoring of an endangered marine vertebrate using precise, non-invasive morphometrics:** Biological Conservation, 2020-02 | journal-article. DOI: 10.1016/j.j.blocon.2019.108402. Part of ISSN: 0006-3207. Source: Aleks Terauds (Preferred source).
  - A snapshot of biodiversity protection in Antarctica:** Nature Communications, 2019-12 | journal-article. DOI: 10.1038/s41467-019-08915-6. Part of ISSN: 2041-1723. Source: Aleks Terauds (Preferred source).
  - Genome-wide SNP data reveal improved evidence for Antarctic glacial refugia and dispersal of terrestrial invertebrates:** Molecular Ecology, 2019-11 | journal-article. DOI: 10.1111/mec.15269.

At the bottom of the page, there are two analytics widgets:

- Website: [researchdata.edu.au](https://researchdata.edu.au)
- website, 0 unique keywords
- 56 downloads, 49 unique downloads

# DIGITAL EARTH: ANTARCTICA

- Stacks of satellite imagery (Landsat, Sentinel) that can be compared over long periods of time
- Monitor changes in land cover (snow, ice), and generate statistics such as frequency of cover, or change trend
- It may also be possible to incorporate radar information to study evolution of crevasses
- Improved communications
- Expanding AAD's remote sensing and imaging capabilities (including drone imagery analysis)





Questions?



© 2011