FAIR data

Science to Service

Luke Marsden

Meteorologisk institutt

UNIS
The University Centre in Svalbard

the Nansen LEGACY

YouTube

ARCTIC PASSION

SIOS
SVALBARD INTEGRATED ARCTIC EARTH OBSERVING SYSTEM


Image by Christian Morel, http://www.c
280 people
10 institutions
the Nansen LEGACY
Multiple disciplines
All data must be FAIR!

"It is one thing to collect and publish all these data, now we need to make sure we reuse the data to support our future research.

--Paraphrasing Jørgen Berge, University of Tromsø
2023, Nansen Legacy symposium

Find	Findability
Access	Accessibility
Understand	Interoperability
Re-use	Reusability


Download all


Science
Service

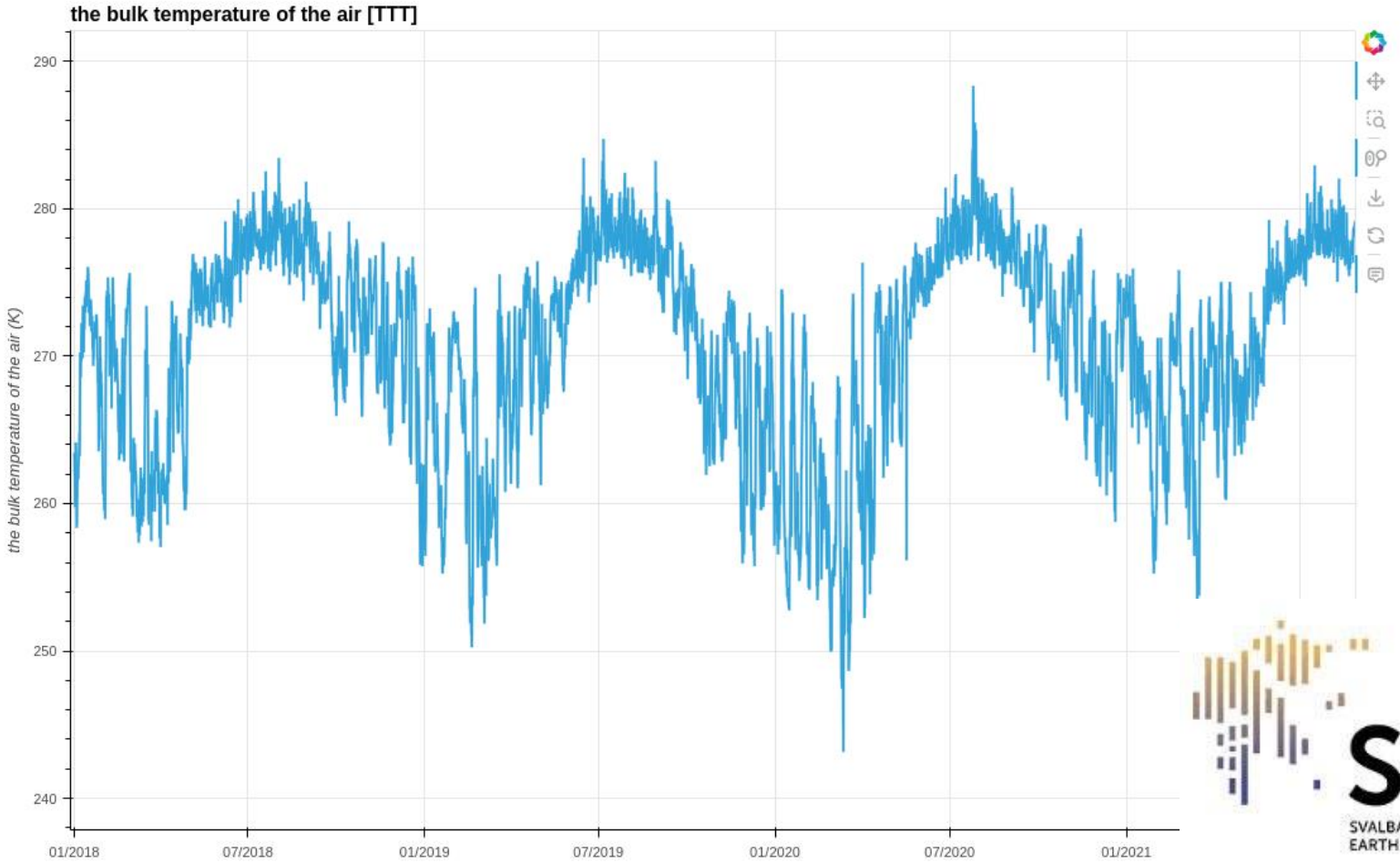
Timeseries Plot

Data Variable

the bulk temperature of the air [TTT]

Export

Metadata



et of the

ergen,

Mea

Stat

SPUB

THE F

Institut

Last me

https

Tempo

Start

End d

Show

Dataset

Data a

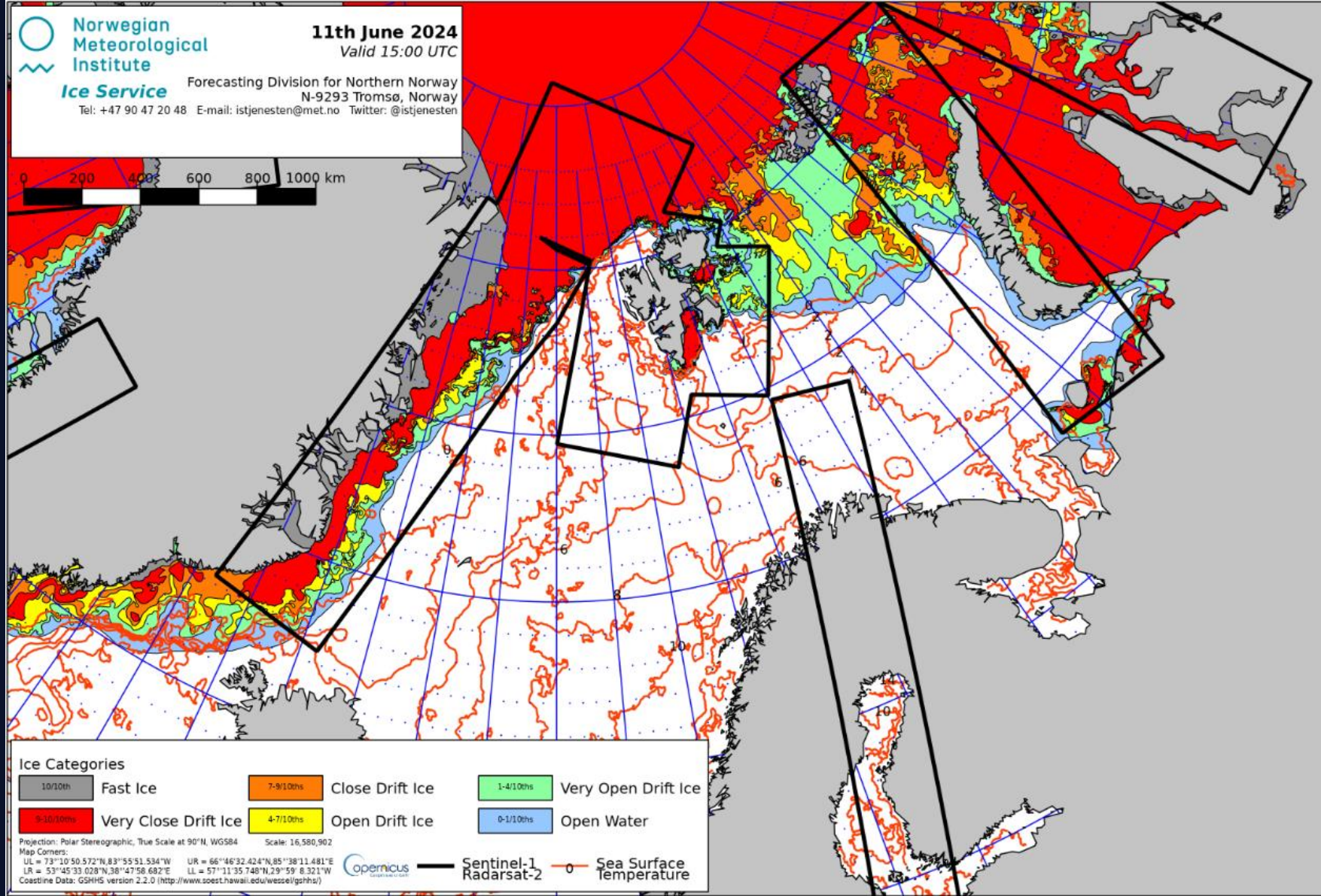
Add

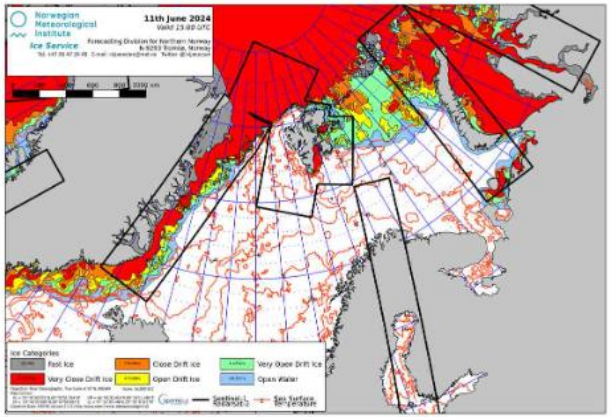
SIOS

SVALBARD INTEGRATED ARCTIC EARTH OBSERVING SYSTEM

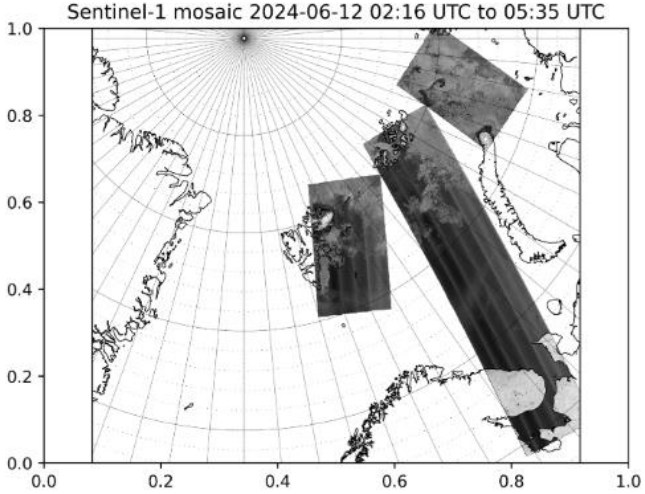
Cryo

https://cryo.met.no

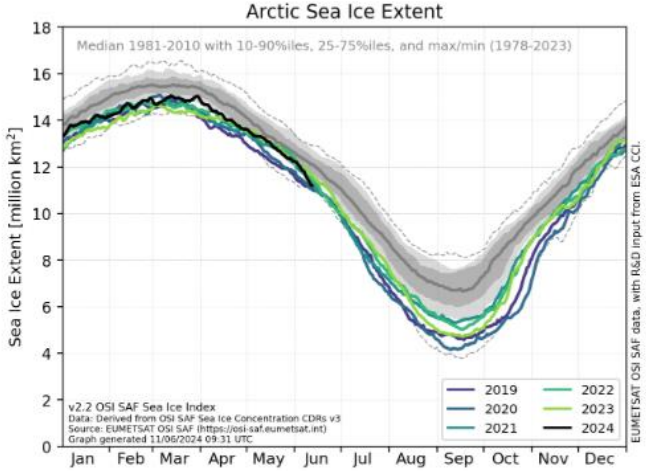




Get latest ice charts

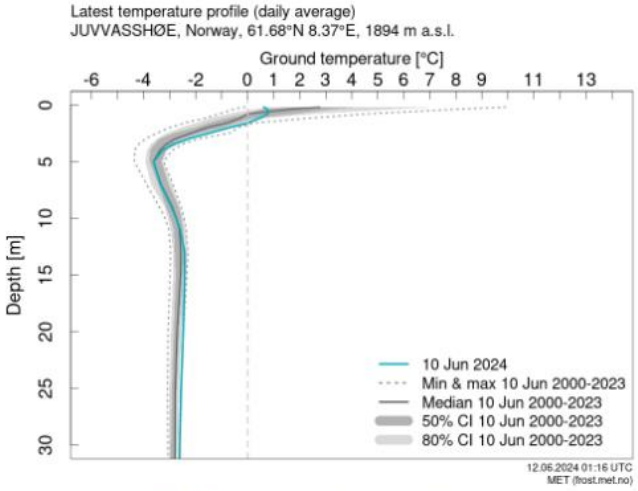


Get latest mosaic



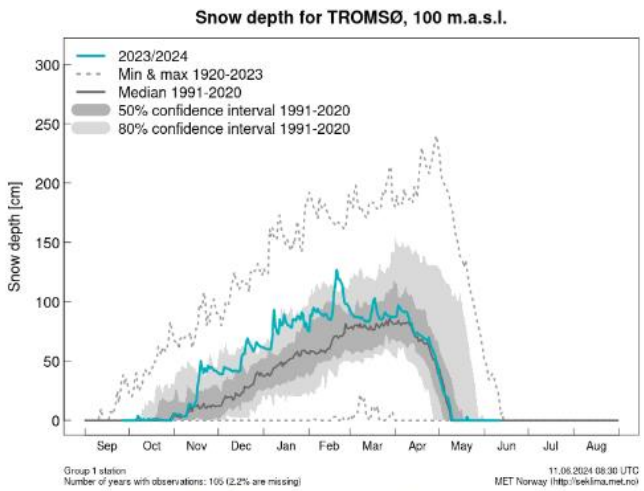
See latest climate trackers

Permafrost Monitoring



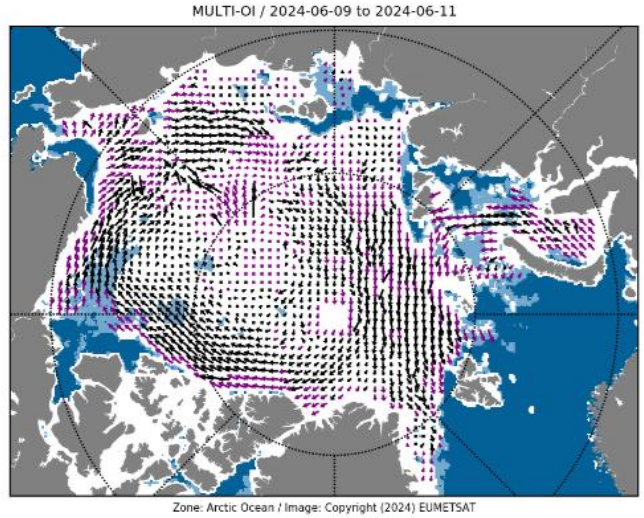
Get latest permafrost products

Snow Monitoring



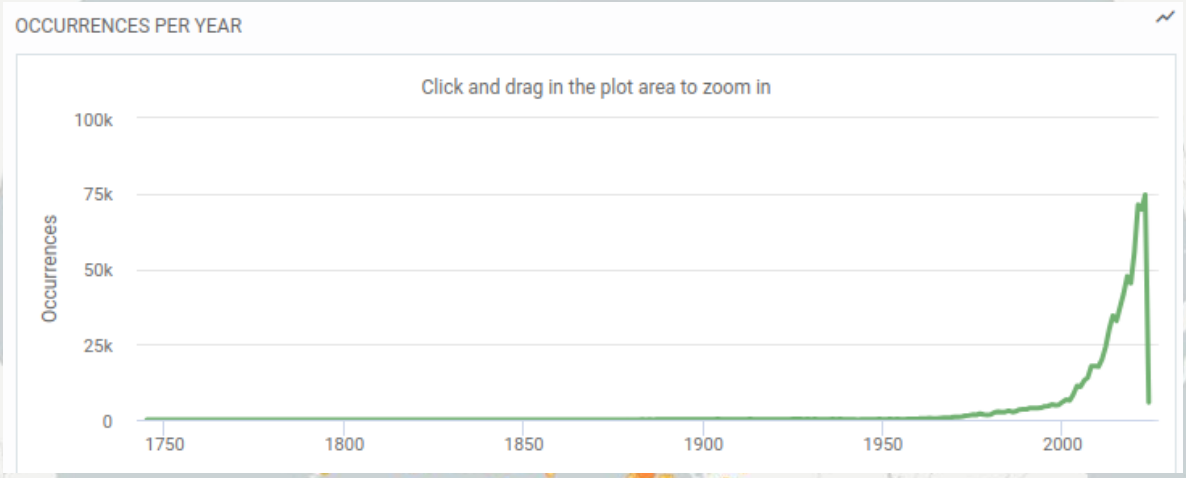
Get latest snow products

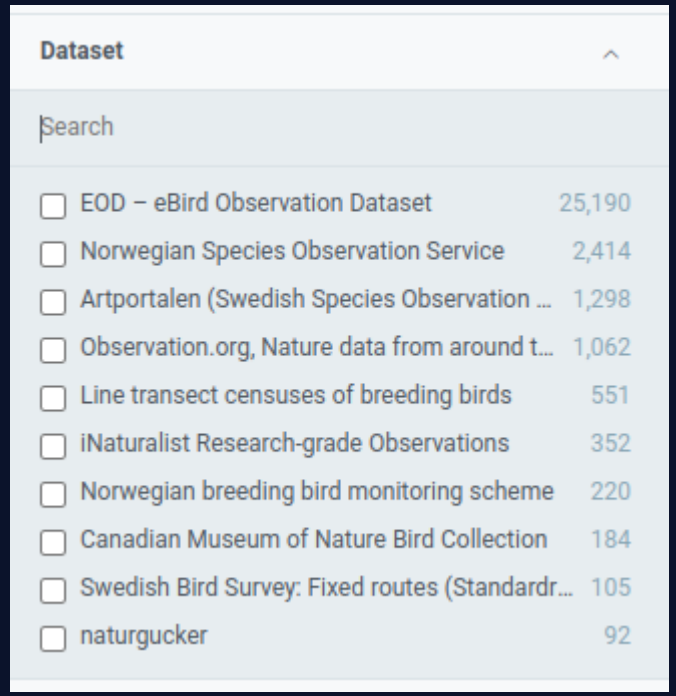
Sea Ice Monitoring



Get latest global sea ice products

Observations in June
Snow bunting





picious coordinates. Hide them

GBIF
Global Biodiversity Information Facility

Plectrophenax nivalis (Linnaeus, 1758)
observed in Svalbard and Jan Mayen
by dougiewainwright (licensed under
http://creativecommons.org/licenses/by-nc/4.0/)

seek

by iNaturalist

Get outside, explore, and learn about the nature all around you!

GET IT ON Google Play

Download on the App Store

@seekbyinat

DESTINATION EARTH

A DIGITAL REPLICA OF OUR PLANET

Destination Earth (**DestinE**) aims to develop a highly accurate digital model of Earth to monitor the effects of natural and human activity on our planet, anticipate extreme events and adapt policies to climate-related challenges.

European Union
ECMWF
esa
EUMETSAT

Consistency and standardisation

FAIR data

FAIR data centres

FAIR data

Findability

Accessibility

Interoperability

Reusability


Data availab
Data will be m
request.
Data Availability Statement
The original contributions presented in t
the article/Supplementary Material, further inquiries
can be directed to the corresponding a

Publish data with a data centre

But there are hundreds of data centres, Luke, which one should we choose!?

Search

Destination/property name:

Longyearbyen

Check-in date

Mon, Mar 27

Check-out date

Fri, Mar 31

4-night stay

2 adults · 0 children · 1 room

I'm traveling for work

Search

Longyearbyen: 5 properties found

Show on map

Sort by: Our Top Picks

Commission paid and other benefits may affect an accommodation's ranking. Learn more.

Filter by:

Your Budget (per night)

Set your own budget

- NOK 1,000 – NOK 1,500 2
- NOK 1,500 – NOK 2,000 2
- NOK 2,000 + 4

Popular Filters

- Breakfast Included 4
- 4 stars 2
- Couples' massage 1
- Less than 1 km 3
  Distance from center of Longyearbyen
- Live music/Performance 3
- Bingo 1
- Racquetball 1
- Hand massage 1

Radisson Blu Polar Hotel, Spitsbergen ★★★★

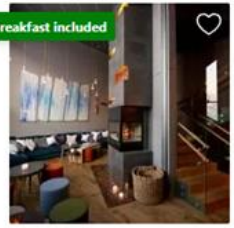
Longyearbyen | Show on map | 300 m from center

Standard Guest Room
2 twin beds
Breakfast included
FREE cancellation • No prepayment needed
You can cancel later, so lock in this great price today!

Very Good 8.1
776 reviews
Location 9.3

4 nights, 2 adults
NOK 15,580
Includes taxes and fees

See availability >



Svalbard Hotell | Polfareren ★★★★

Longyearbyen | Show on map | 250 m from center

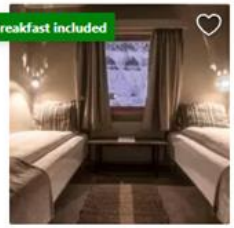
Travel Sustainable Level 1

Standard Double or Twin Room
Beds: 1 double or 2 twins
Breakfast included
Free cancellation
You can cancel later, so lock in this great price today!
Only 7 rooms left at this price on our site

Excellent 8.8
434 reviews
Location 9.5

4 nights, 2 adults
NOK 16,845
Includes taxes and fees

See availability >



Coal Miners' Cabins

Longyearbyen | Show on map | 2.2 km from center

Economy Twin Room with Shared Bathroom
2 twin beds
Breakfast included
Only 3 rooms left at this price on our site

Very Good 8.2
161 reviews

4 nights, 2 adults
~~NOK 6,748~~ NOK 6,302
Includes taxes and fees

See availability >

Data centre examples

- Good
  - https://www.gbif.org/
  - https://adc.met.no/
- Less good
  - https://zenodo.org/
  - https://datadryad.org

trivago

Mine favoritter | Logg inn | Meny | NOK

Longyearbyen | Ankomst ma., 27.03.23 | Avreise fr., 31.03.23 | 1 rom 2 gjester | Søk

Pris per natt kr 0 - kr 5 400+ | Type overnatting: Alle | Hotell | Hus/leilighet | Vurdering: Alle | Beliggenhet | Flere filtre: Velg

Sorter etter: Våre anbefalinger

Overnattingssteder funnet: 5 Bookingsider søkt på: 178

Coal Miners' Cabin

Hostell/vandrerhjem

2.5 km til Sentrum

8.2 Veldig bra (700 anmeldelser)

22 % lavere enn vanlig

Booking.com
Frokost inkludert
4 netter: kr 6 302
kr 1 575
Vis tilbud

opodo
kr 1 598

Vår laveste pris:
kr 1 575 Booking.com

Svalbard Hotell | Polfareren

★★★★ Hotell

0.4 km til Sentrum

9.1 Fantastisk (665 anmeldelser)

Booking.com
Gratis avbestilling · Frokost inkludert
4 netter: kr 16 845
kr 4 211
Vis tilbud

opodo
kr 3 790

Vår laveste pris:
kr 3 722 Stayforlong

Gjestehuset 102

★★ Hostell/vandrerhjem

2.6 km til Sentrum

8.3 Veldig bra (313 anmeldelser)

Booking.com
Frokost inkludert
4 netter: kr 5 969
kr 1 492
Vis tilbud

opodo
kr 1 562

Vår laveste pris:
kr 1 492 Booking.com

Radisson Blu Polar Hotel Spitsbergen

Hotell

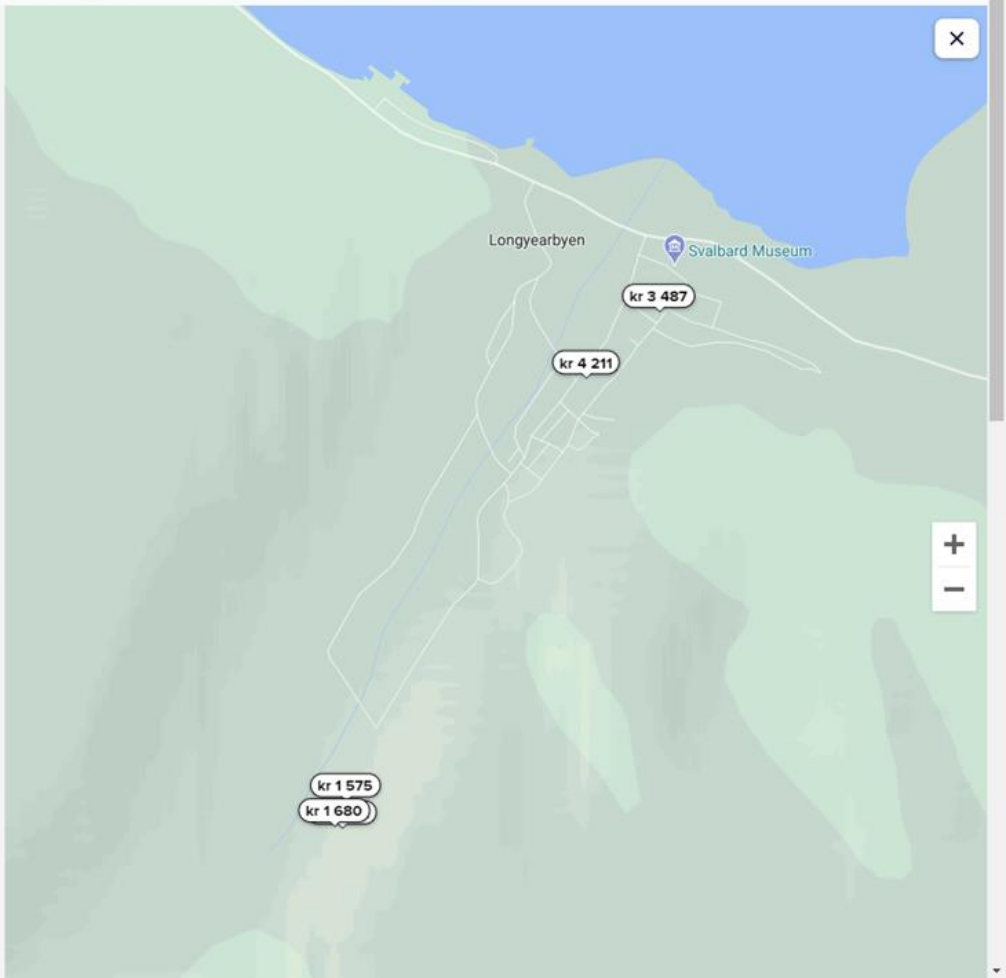
0.0 km til Sentrum

8.2 Veldig bra (1031 anmeldelser)

Trip.com
Gratis avbestilling · Frokost inkludert
4 netter: kr 13 947
kr 3 487
Vis tilbud

Radisson Hotels

Vår laveste pris:



Data access portals

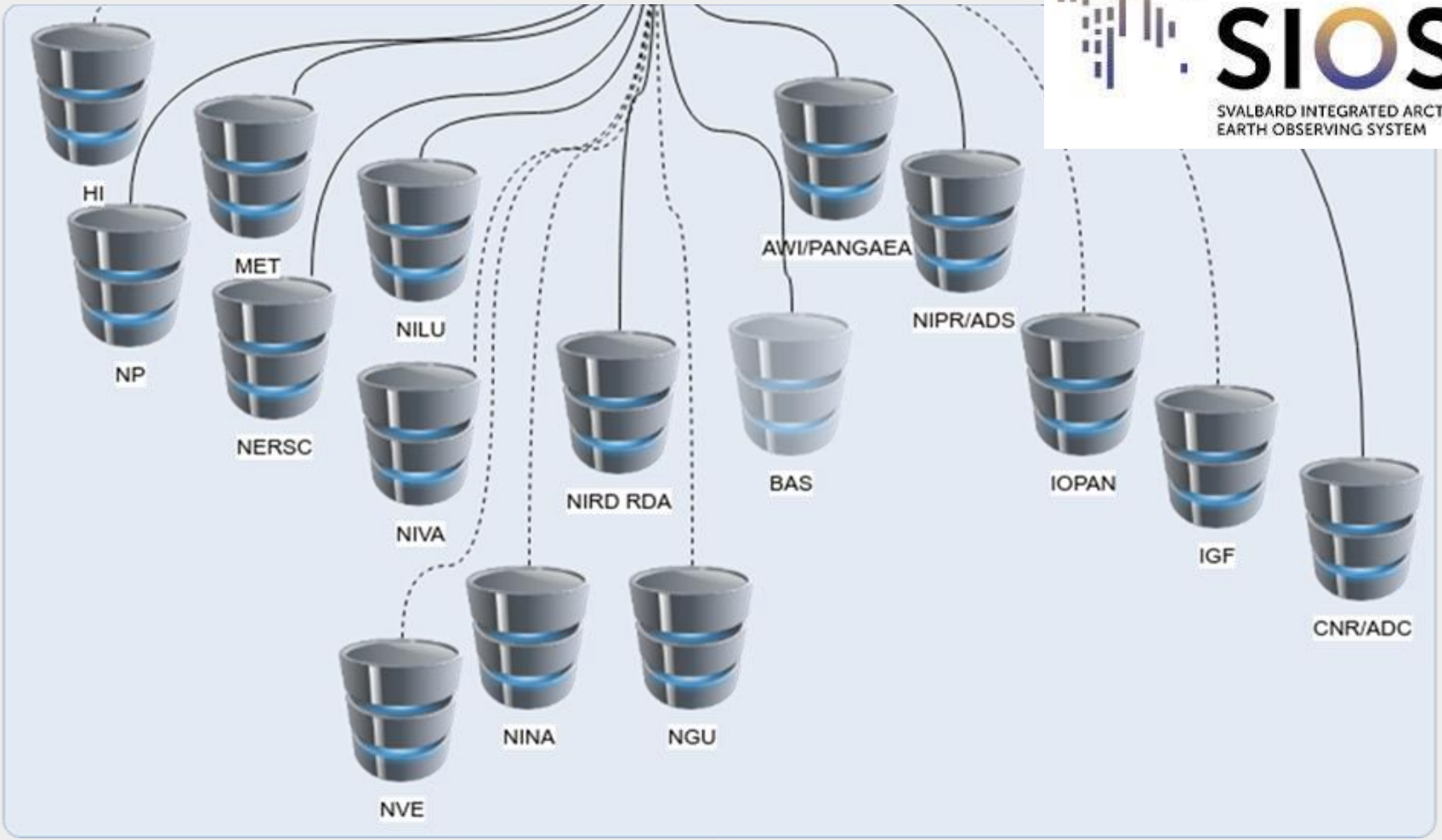
https://sios-svalbard.org/metsis/search?f%5B0%5D=dataset_level%3ALevel-1

https://data.arcticobserving.org/

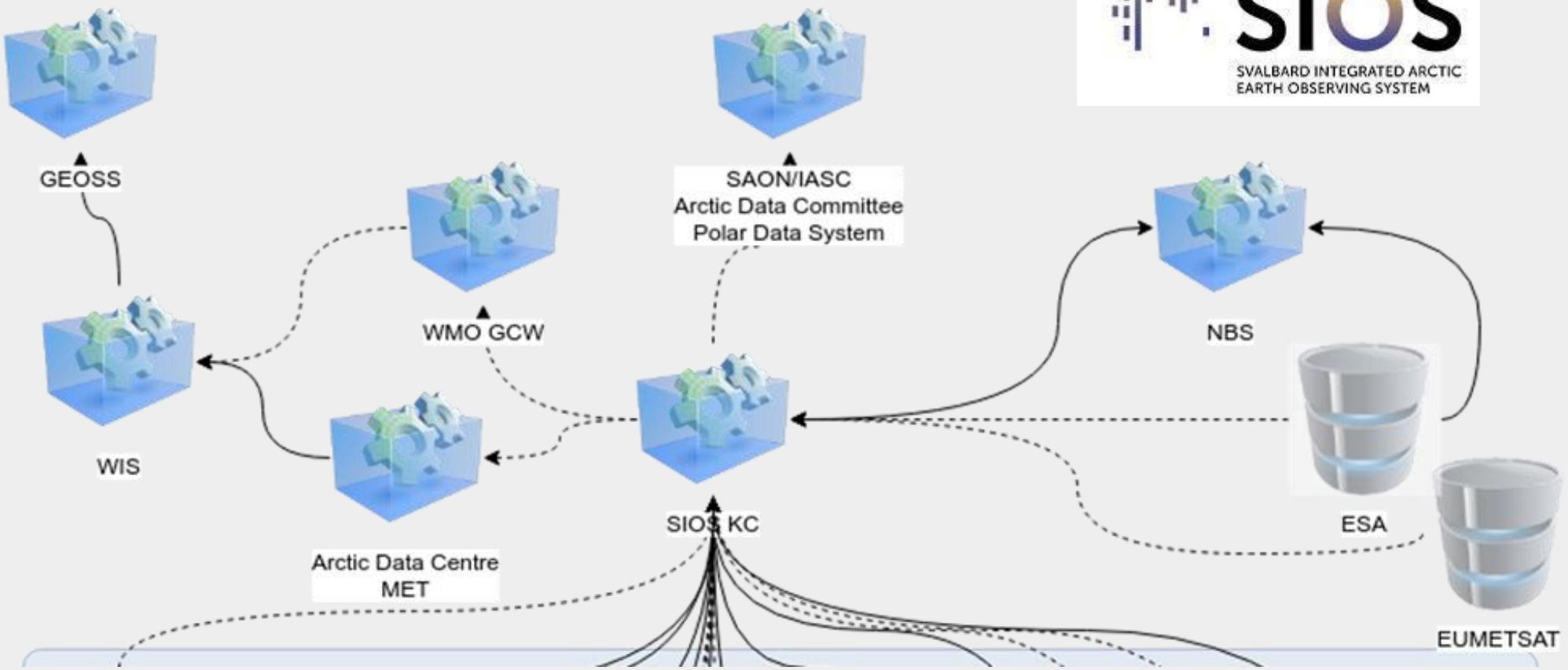
https://search.polder.info/

Tip: Check which data centres contribute to these!

SIOS data access portal



SIOS data access portal




Science
Service

un-FAIR data

Date	Time	Temperature	Salinity	X

FAIR data

Standardised structure

Include metadata

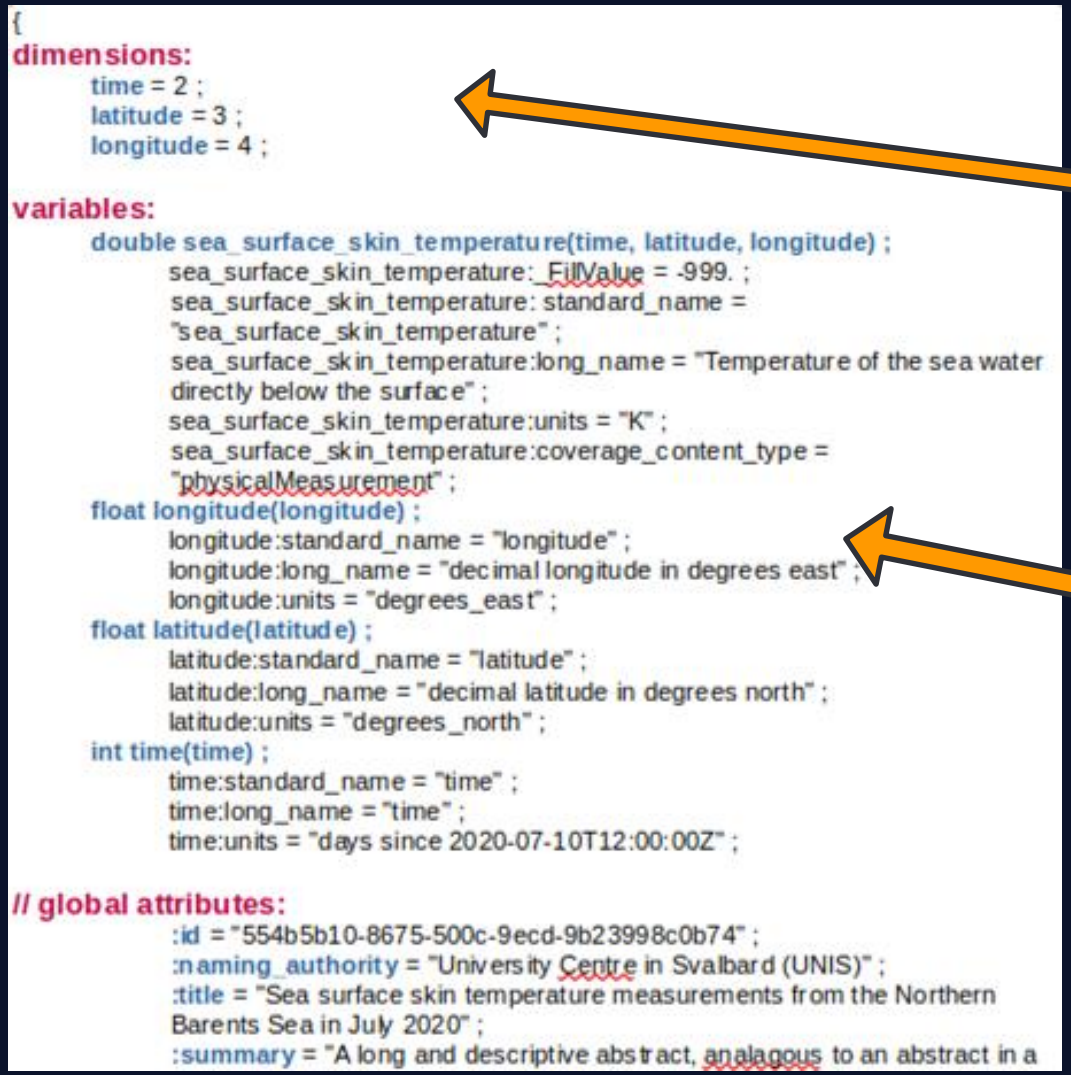
Darwin Core Archive

Controlled vocabularies

CF-NetCDF

Machine readable

Software independent



Dimensions:

- 3 points of latitude
- 4 points of longitude
- 2 points time

Variables:

- Variable ‘lat’ has dimension lat
- Variable ‘…temperature…’ has 3 dimensions
- Variable attributes

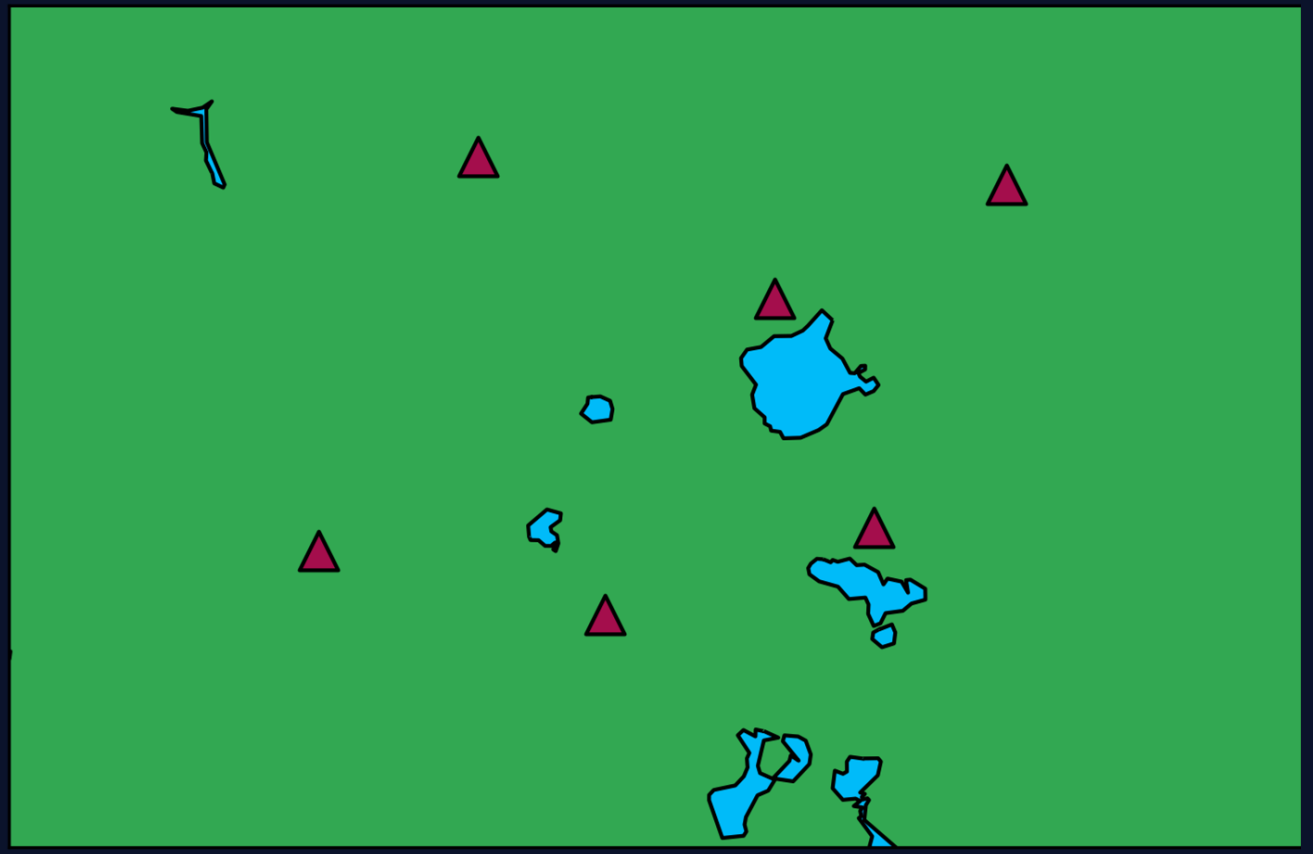
// global attributes:

:id = "554b5b10-8675-500c-9ecd-9b23998c0b74" ;
:naming_authority = "University Centre in Svalbard (UNIS)" ;
:title = "Sea surface skin temperature measurements from the Northern Barents Sea in July 2020" ;
:summary = "A long and descriptive abstract, analagous to an abstract in a journal article, that describes the data" ;
:keywords = "Earth Science > Oceans > Ocean Temperature > Sea Surface Temperature > Sea Surface Skin Temperature" ;
:keywords_vocabulary = "GCMD" ;
:geospatial_lat_min = 76LL ;
:geospatial_lat_max = 78LL ;
:geospatial_lon_min = 28LL ;
:geospatial_lon_max = 31LL ;
:time_coverage_start = "2020-07-10T12:00:00Z" ;
:time_coverage_end = "2020-07-19T12:00:00Z" ;
:Conventions = "ACDD-1.3, , CF-1.8" ;
:history = "File created at 2022-01-10T14:21:58 using the xarray library in Python" ;
:source = "Sea surface temperature measurements recorded by....okay, I made them up!" ;
:processing_level = "raw" ;
:date_created = "2022-01-10T14:21:58Z" ;
:creator_type = "person; person" ;
:creator_institution = "The University Centre in Svalbard, Norway; The University Centre in Svalbard, Norway" ;
:creator_name = "Luke Marsden; John Doe" ;
:creator_email = "lukem@unis.no; jodoe@unis.no" ;
:creator_url = "https://www.unis.no/staff/luke-marsden/; https://www.unis.no/staff/john-doe/" ;
:institution = "The University Centre in Svalbard (UNIS)" ;
:publisher_name = "Norwegian Meteorological Institute - Arctic Data Centre" ;
:publisher_email = "adc-suppoort@met.no" ;
:publisher_url = "https://adc.met.no/" ;
:publisher_type = "institution" ;
:project = "The Nansen Legacy (RCN # 276730)" ;
:license = "https://creativecommons.org/licenses/by/4.0/" ;
:metadata_link = "" ;
:metadata_catalogue = "https://sios-svalbard.org/aen/tools, event ID = a5ec89da-f45e-11eb-b6ae-0f3712703d75" ;
:acknowledgements = "Funded by the Research Council of Norway. John Smith was involved in collecting the data" ;

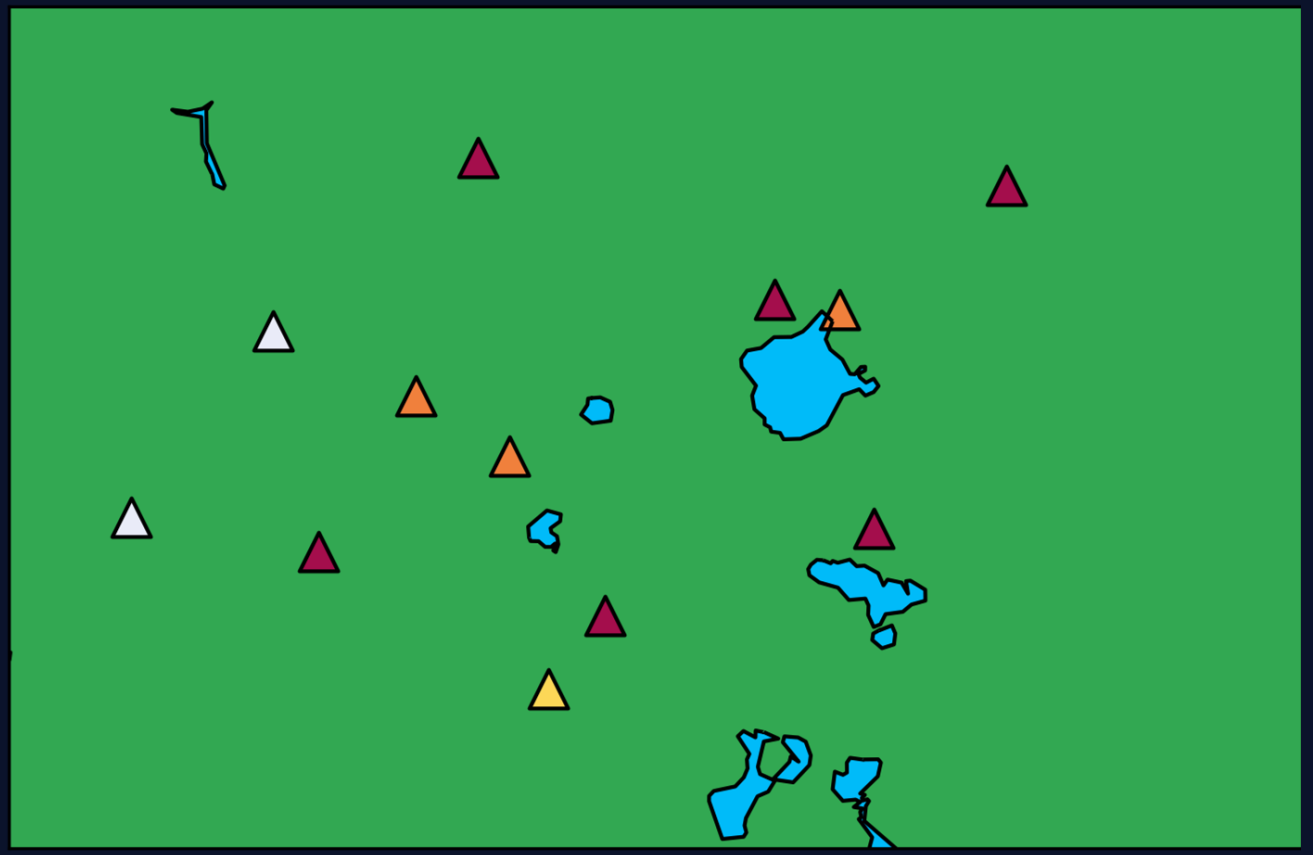
Conventions used:
CF
ACDD

More tomorrow!

Granularity



Granularity



Granularity

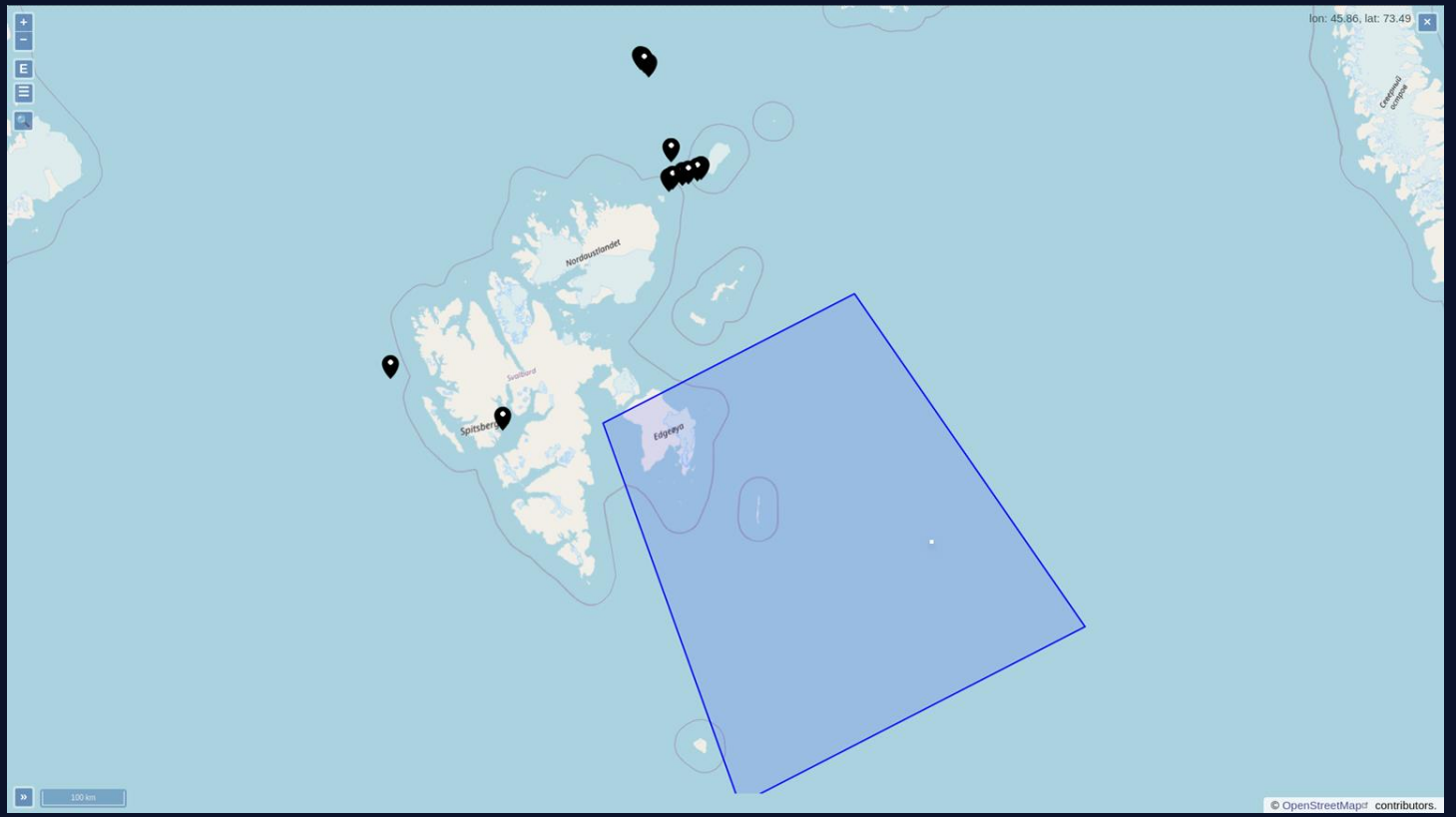
Technical documentation

Granularity Perspectives Document

Version 0.0, 2024-01-11: Draft

Advantages of finer granularity

- Metadata



Advantages of finer granularity

- Metadata
- Simpler files

```
<xarray.Dataset>
Dimensions:                (depth: 3)
Coordinates:
  * depth                  (depth) int64 10 20 30
Data variables:
    sea_water_temperature  (depth) float64 21.42 21.21 20.98
```

```
<xarray.Dataset>
Dimensions:                (time: 4, depth: 3, latitude: 4, longitude: 4)
Coordinates:
  * latitude               (latitude) float64 78.51 79.28 79.98 80.42
  * longitude              (longitude) float64 30.42 30.36 30.5 30.42
  * depth                  (depth) int64 10 20 30
  * time                   (time) int64 0 1 2 3
Data variables:
    sea_water_temperature  (time, depth) float64 21.42 21.21 ... 21.01 19.99
```

Advantages of finer granularity

- Metadata
- Simpler files
- Less empty space

```
array([[21.42,   nan,   nan, 21.21,   nan,   nan,   nan,   nan, 20.98,
          nan,   nan,   nan,   nan,   nan],
       [  nan, 22.08,   nan,   nan,   nan,   nan, 21.56,   nan,   nan,
          nan, 20.42,   nan, 19.23, 18.53],
       [  nan, 22.42,   nan,   nan,   nan, 21.21,   nan, 20.12,   nan,
        19.45,   nan, 18.72,   nan, 16.99],
       [  nan,   nan, 21.84,   nan, 21.49,   nan,   nan,   nan,   nan,
          nan,   nan,   nan,   nan,   nan]])
```

```
<xarray.Dataset>
Dimensions:                 (time: 4, depth: 14)
Coordinates:
  * depth                   (depth) int64 5 10 12 20 24 25 ... 50 60 70 80 90 100
  * time                    (time) int64 0 1 2 3
Data variables:
    sea_water_temperature   (time, depth) float64 21.42 nan nan ... nan nan nan
    latitude                (time) float64 78.51 79.28 79.98 80.42
    longitude               (time) float64 30.42 30.36 30.5 30.42
```

Advantages of finer granularity

- Metadata
- Simpler files
- Less empty space
- Easier to build services upon

Parent & child

IWIN: The Isfjorden Weather Information Network

By laraf | Wed, 03/15/2023 - 14:47

DOI: https://doi.org/10.21343/e[illegible]w846

Citation

Frank, L., Jonassen, M. O., & Remes, T. (2023).

License: CC BY

Related Datasets

HasPart

https://doi.org/10.21343/bm2t-r768
https://doi.org/10.21343/n2wv-0t61
https://doi.org/10.21343/r091-2n20
https://doi.org/10.21343/5x4s-hx67

Accessing many files in one go (THREDDs)

```
]:
base_url = 'https://opendap1.nodc.no/opendap/phys
catalog_url = base_url + '/catalog.xml'
catalog = TDSCatalog(catalog_url)

for dataset in catalog.datasets:
    profile_url = base_url + '/' + dataset
    xrds = xr.open_dataset(profile_url)
    print(xrds.attrs['time_coverage_start'])
```

```
2021-07-12T19:05:04Z
2021-07-14T23:59:51Z
2021-07-15T02:01:30Z
2021-07-15T03:35:31Z
2021-07-15T04:45:26Z
2021-07-15T05:55:33Z
2021-07-15T07:25:12Z
2021-07-15T08:44:11Z
2021-07-16T17:22:02Z
```

OPeNDAP

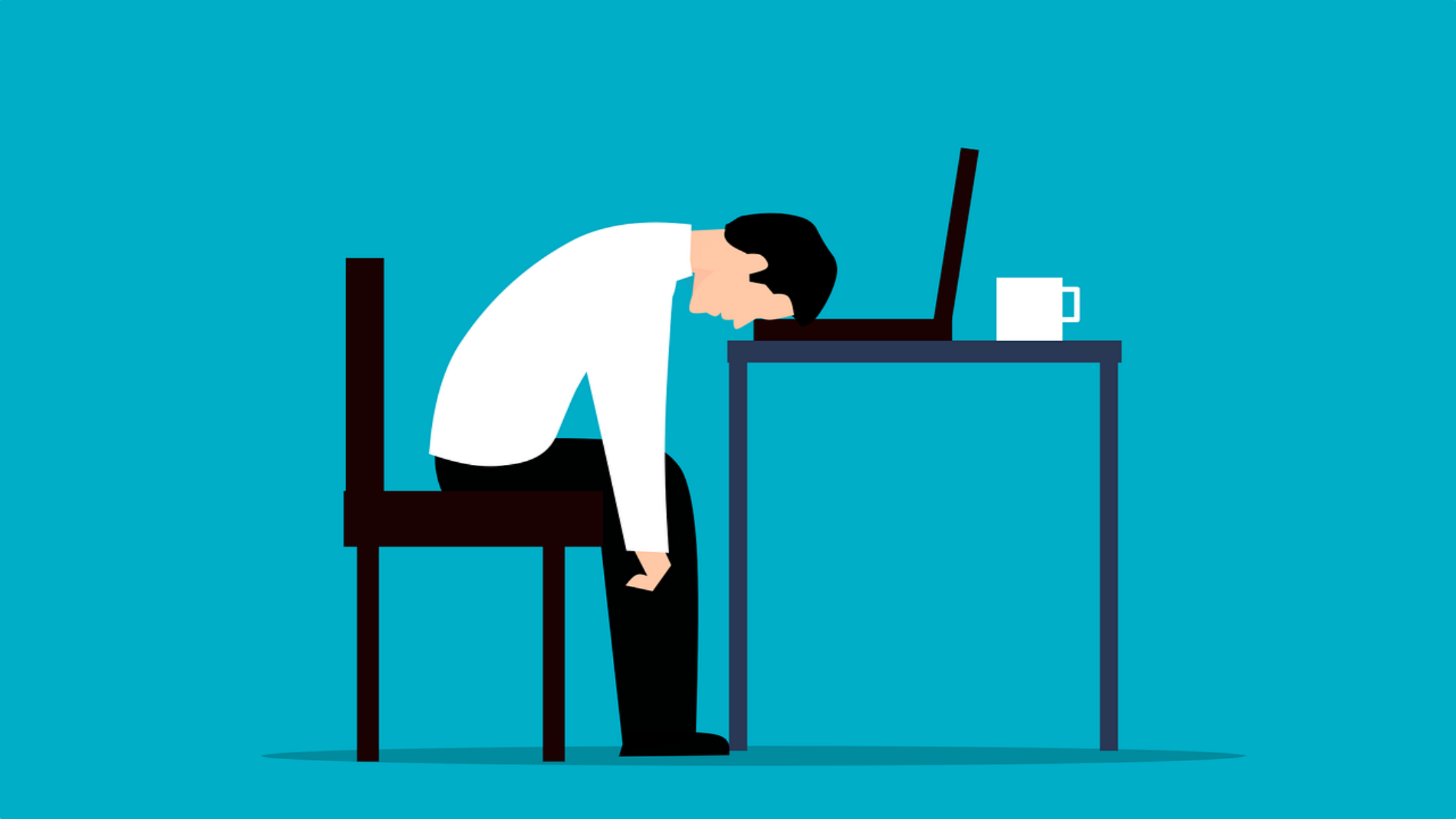
Contents of /physics/point/cruise/nansen_legacy-single_profile/NMDC_

Name

CTD_station_ISG_SVR1_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG02_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG02_1_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG02_2_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG02_3_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG02_4_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG02_5_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG03_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG05_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG05_01_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG05_02_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG06_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG06_01_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG06_02_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG07_01_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG08_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG08_1_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG09_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG09_1_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG10_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc
CTD_station_NLEG10_1_-_Nansen_Legacy_Cruise_-_2021_Joint_Cruise_2-1.nc

The FAIR data revolution

How to make it happen…



Incentives

Regulations

Creating incentives

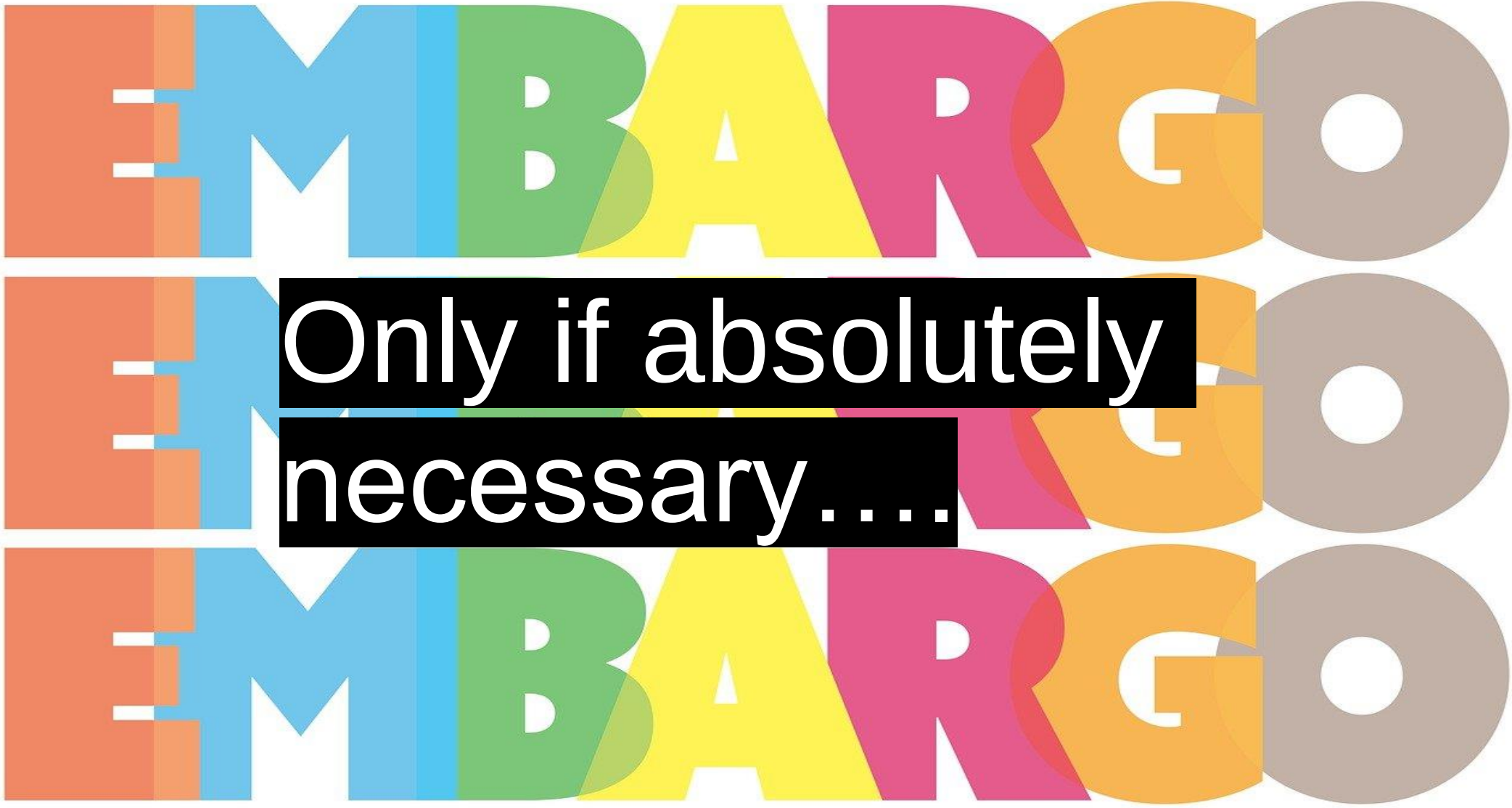
44 OCCURRENCES | 5 CITATIONS

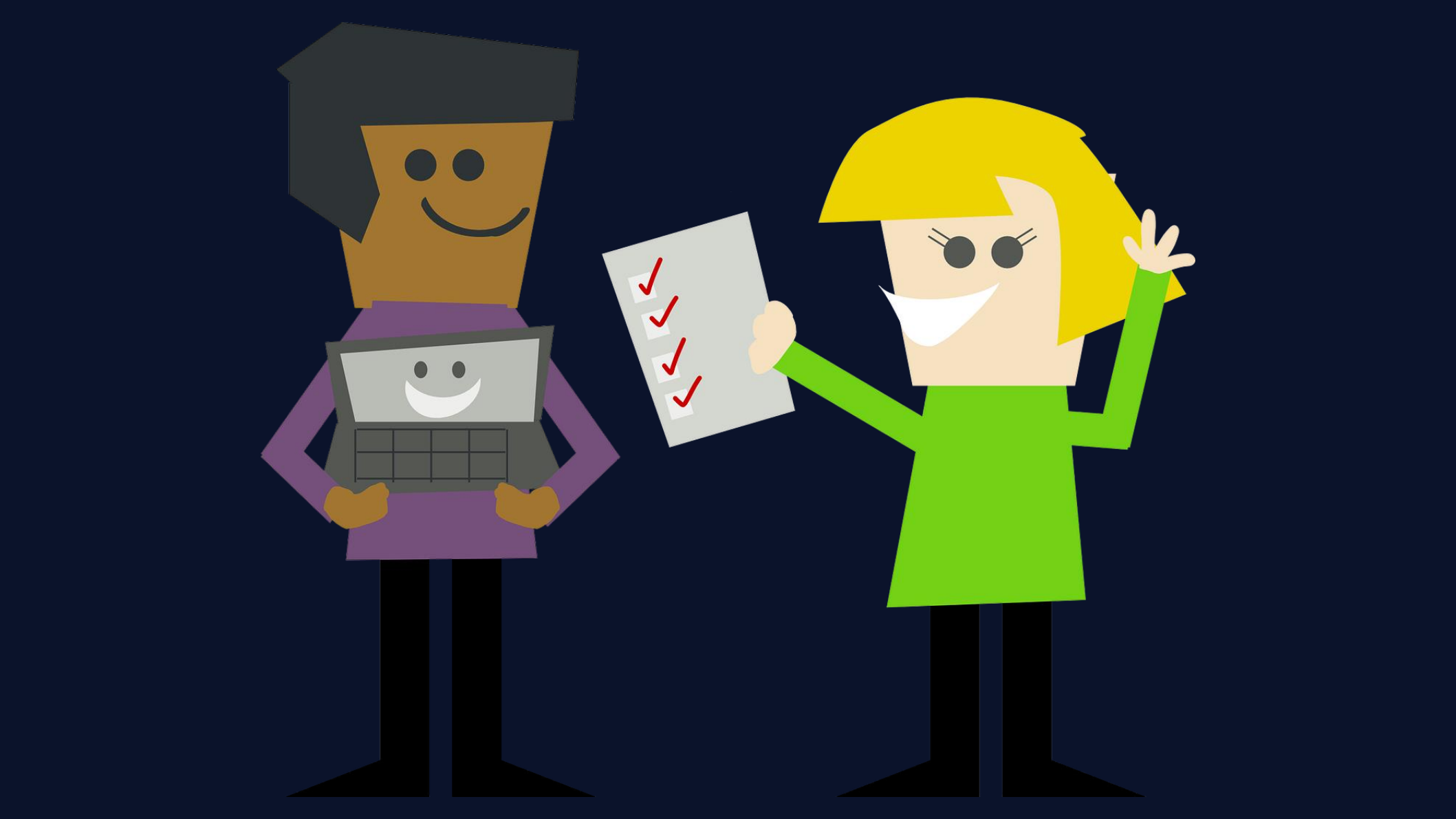
Citation

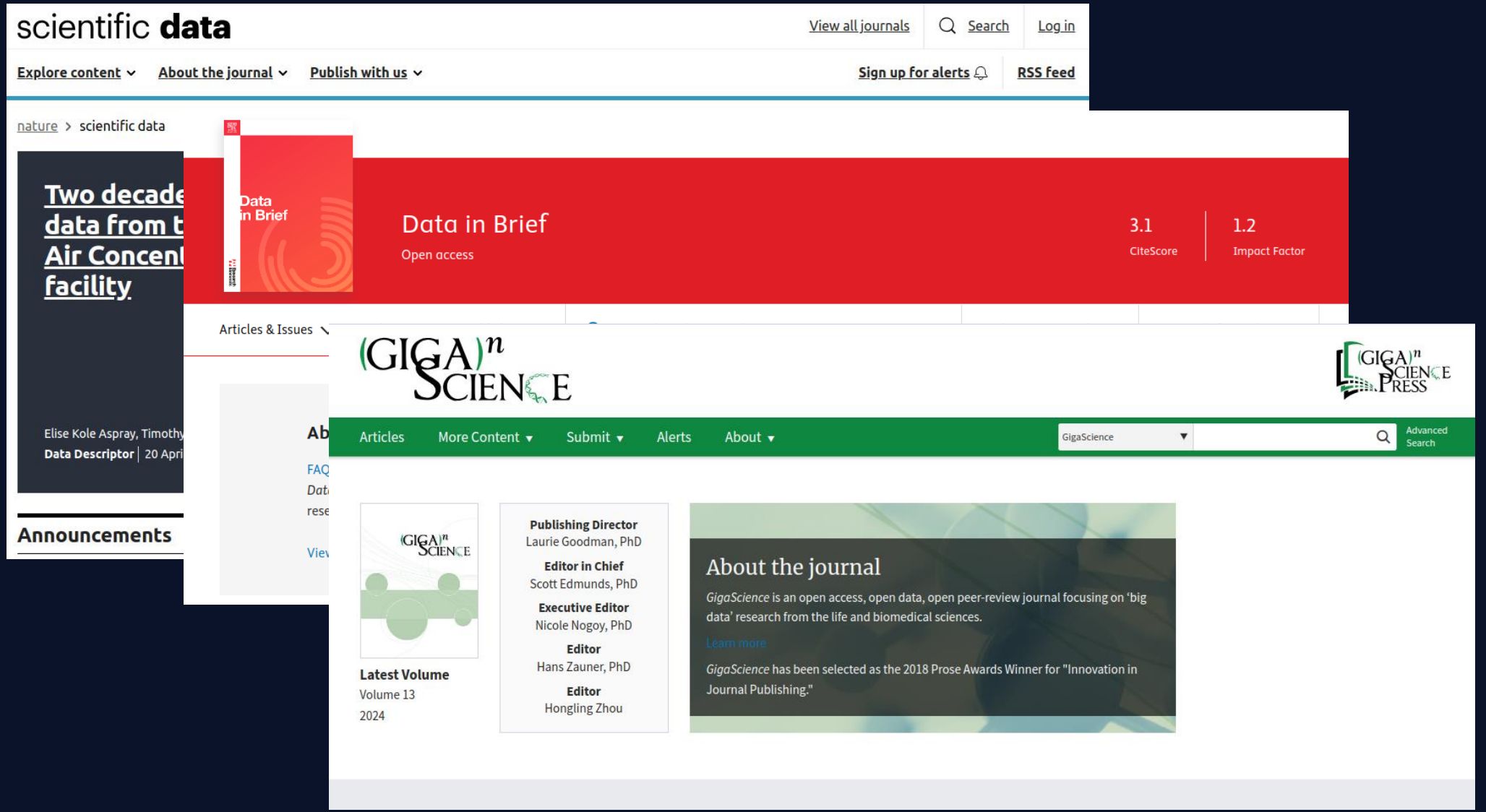
Aakre M, Berntsen S T, Broderstad S L, Jensen S S, Frisholm J M (2023). UNISAB202-2023-benthos. Version 1.1. The University Centre in Svalbard. Occurrence dataset. https://ipt.gbif.no/resource?r=unisab202-2023-benthos&v=1.1 https://doi.org/10.15468/2qdh2f accessed via GBIF.org on 2024-06-17.

Creating incentives

License	Type of use	Symbols
Attribution (BY)	You must credit the creator, the title and the license the work is under.	CC BY
Non Commercial (NC)	The work cannot be used for commercial purposes.	CC BY NC
No Derivatives (ND)	The work can only be used exactly how it is. The work cannot be adapted or modified in any way.	CC BY ND
Share Alike (SA)	Any new material produced using the work must be made available under the same license as the original work.	CC BY SA
Public Domain CC0	The copyright holder has voluntarily waived their rights and the work can be freely used	PUBLIC DOMAIN


EMBARGO
EMBARGO
EMBARGO
Only if absolutely necessary….



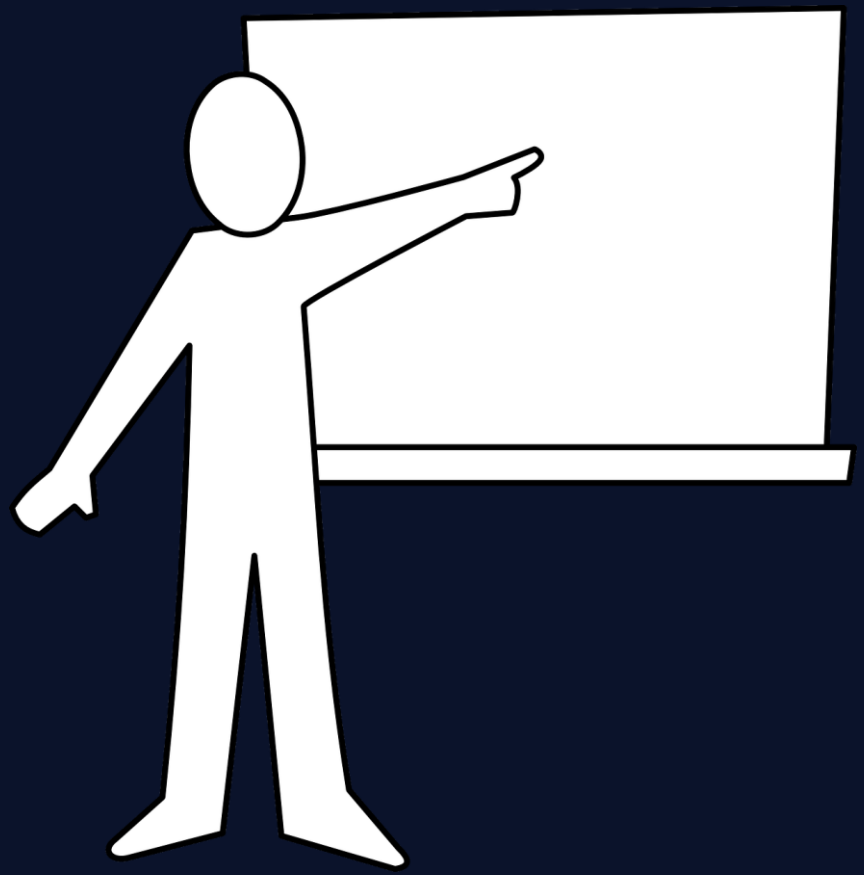

scientific data
View all journals
Search
Log in
Explore content
About the journal
Publish with us
Sign up for alerts
RSS feed
nature > scientific data
Two decade
data from t
Air Concent
facility
Elise Kole Aspray, Timothy
Data Descriptor | 20 Apri
Announcements
Data in Brief
Data in Brief
Open access
3.1
CiteScore
1.2
Impact Factor
Articles & Issues
(GIGA)n SCIENCE
(GIGA)n SCIENCE PRESS
Articles
More Content
Submit
Alerts
About
GigaScience
Advanced Search
Latest Volume
Volume 13
2024
Publishing Director
Laurie Goodman, PhD
Editor in Chief
Scott Edmunds, PhD
Executive Editor
Nicole Nogoy, PhD
Editor
Hans Zauner, PhD
Editor
Hongling Zhou
About the journal
GigaScience is an open access, open data, open peer-review journal focusing on 'big data' research from the life and biomedical sciences.
Learn more
GigaScience has been selected as the 2018 Prose Awards Winner for "Innovation in Journal Publishing."

Data policies + Data management plans

Institutions

Projects

Funder requirements



My plan for next 2 days

Now: Q&A

Session 2: SIOS data access portal

Sessions 4 & 5 (tomorrow):
Working with CF-NetCDF files in Python or R

YouTube

https://www.youtube.com/@LukeDataManager