



# **Data management for observational data at UiB: a chance to strengthen our research environment, increase UiB's credibility and visibility, and secure our scientific heritage**

**Benjamin Pfeil**

**Bjerknes Climate Data Centre  
Geophysical Institute  
University of Bergen**

# Structure

---

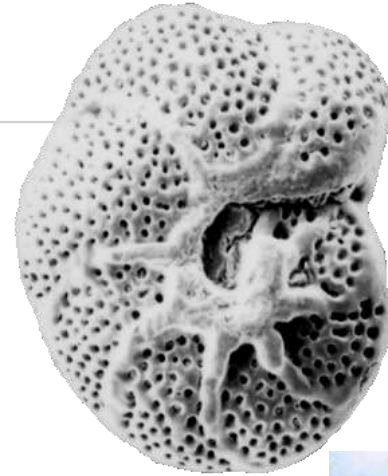
- › What is observational data?
- › Current data management activities at UiB's Faculty of Mathematics and Natural Sciences
- › Formal requirements
- › Bjerknes Climate Data Centre
- › Vision for UiB
- › Showcase RI ICOS
- › Summary



# What is observational data?

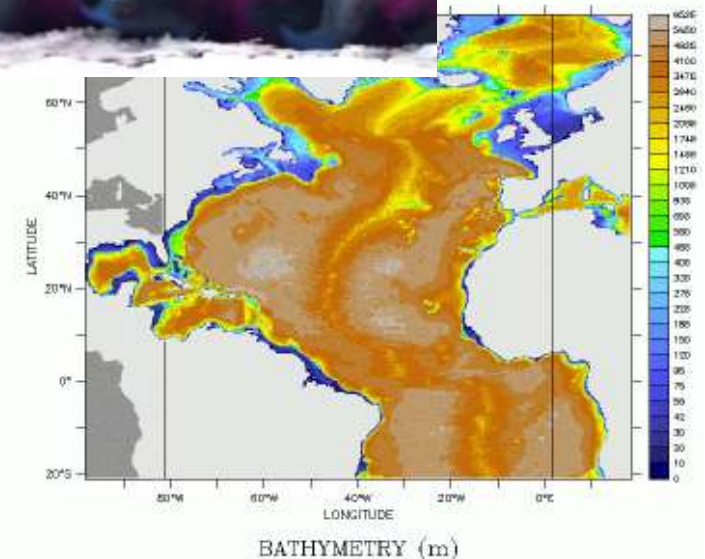
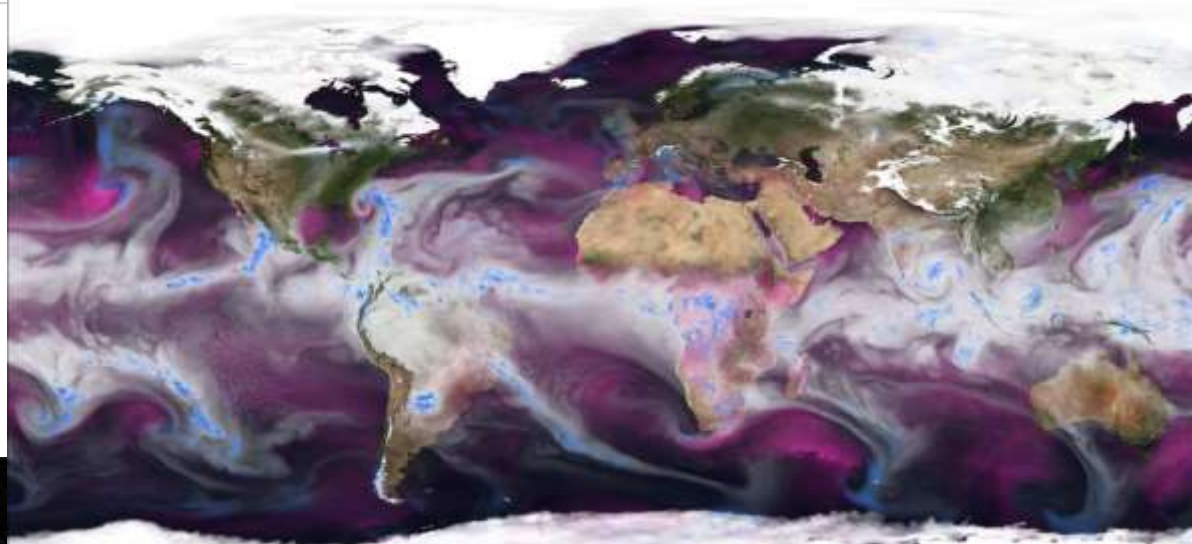


# But also





Within UiB's MatNat we also deal with model output and partly with remote sensing data



## Observational data in the field of natural sciences

---

- › Shows a snapshot of the environment at that time/space
- › Sampling can be very expensive (average of over 100.000 € for one dataset for bio-, geosciences - including costs for expeditions, laboratories, etc)
- › Data is very valueable for future scientific work and has to be archived and made available

# Why do we need data?

---

- › Verification of research results
- › Comparison of results
- › Indication of trends
- › Model input
- › Remote sensing
- › ...

## Data in the scientific community

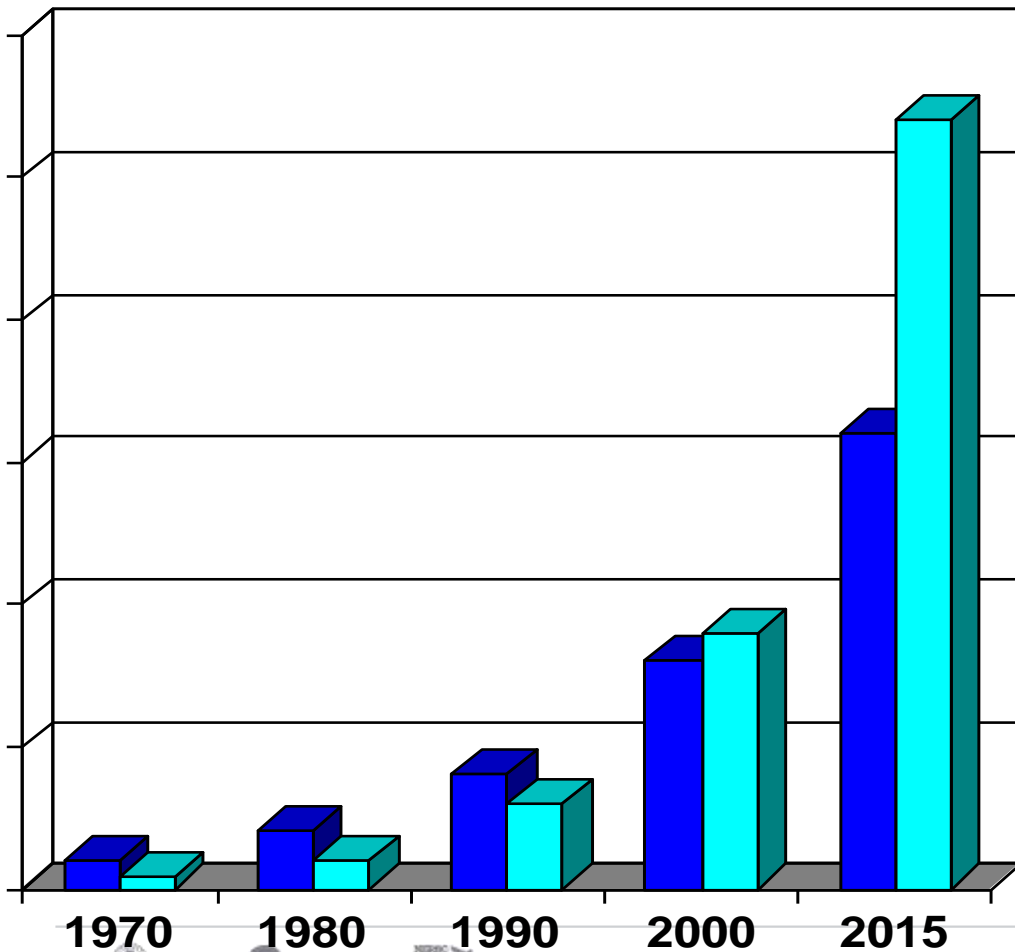
---

- › Scientific instruments and computer simulations create larger and larger amounts of data
- › Due to new measurement procedures (and better precision) are the data volumes are doubling every few years (Big Data)
- › Scientific data has to be archived according to "Good scientific practise in research and scholarship" (*European Science Foundation 2000*)





# Global increase in publications in empirical sciences



newsblog

Nature brings you breaking news from the world of science

News & Comment > News blog Archive > Post

NEWS BLOG

## Global scientific output doubles every nine years

07 May 2014 | 16:46 BST | Posted by Richard Van Noorden | Category: Policy, Publishing

NEWS BLOG













- Publications
- Data

# What are scientific data and how can they be structured?

---

# Data

[Download dataset as tab-delimited text](#) (use the following character encoding: UTF-8: Unicode (PANGAEA default))

1  <b>Event</b>	2  <b>Date/Time</b>	3  <b>Latitude</b>	4  <b>Longitude</b>	5  <b>Depth water [m]</b>	6   <b>Press [dbar]</b>	7   <b>d13C PI [per mil PDB]</b>	8  <b>Lab no</b>
3175MB93/35-43 	1993-08-04T14:41	30.0040	-23.5310	297	299.7	1.694989	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	396	399.6	1.568700	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	544	549.1	1.322923	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	693	699.0	1.151356	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	940	949.5	0.985546	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	1237	1249.8	0.968566	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	1485	1501.1	1.003125	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	1732	1751.6	1.019520	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	1978	2001.8	1.009738	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	2471	2503.8	1.011724	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	2964	3006.5	0.970000	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	3944	4009.1	0.970000	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	4433	4511.7	0.920000	22035
3175MB93/35-43	1993-08-04T14:41	30.0040	-23.5310	5244	5347.0	0.930000	22035
3175MB93/38-46 	1993-08-06T11:24	23.9930	-26.8650	198	200.1	1.932960	22038
3175MB93/38-46	1993-08-06T11:24	23.9930	-26.8650	297	299.7	1.578329	22038
3175MB93/38-46	1993-08-06T11:24	23.9930	-26.8650	397	400.3	1.254143	22038
3175MB93/38-46	1993-08-06T11:24	23.9930	-26.8650	553	558.3	0.950251	22038
3175MB93/38-46	1993-08-06T11:24	23.9930	-26.8650	691	697.3	0.700484	22038
3175MB93/38-46	1993-08-06T11:24	23.9930	-26.8650	845	852.9	0.729354	22038

Citation:

**Olsen, Are; Ninnemann, Ulysses S (2010):** Reconstructed preindustrial d13C of Malcolm Baldrige cruise 3175MB93 on WOCE section AR21. *Bjerknes Centre for Climate Research*, doi:10.1594/PANGAEA.745440



Always quote above citation when using data! You can download the citation in several formats below.

- [RIS Citation](#)
- [BibTeX Citation](#)
- [Text Citation](#)
- [Facebook](#)
- [Twitter](#)
- [Google+](#)
- [Show Map](#)
- [Google Earth](#)

**Related to:** **Olsen, Are; Ninnemann, Ulysses S (2010):** Large d13C Gradients in the Preindustrial North Atlantic Revealed. *Science*, **330(6004)**, 658-659, doi:10.1126/science.1193769 [Q](#)

**Coverage:** *Median Latitude:* 44.528353 \* *Median Longitude:* -21.554647 \* *South-bound Latitude:* 21.007000 \* *West-bound Longitude:* -28.465000 \* *North-bound Latitude:* 62.987000 \* *East-bound Longitude:* -19.915000

*Date/Time Start:* 1993-08-04T14:41:00 \* *Date/Time End:* 1993-08-28T23:26:00

*Minimum DEPTH, water:* 198 m \* *Maximum DEPTH, water:* 5244 m

**Event(s):** **3175MB93/35-43** [Q](#) [📄](#) \* *Latitude:* 30.004000 \* *Longitude:* -23.531000 \* *Date/Time:* 1993-08-04T14:41:00 \* *Campaign:* 3175MB93 [Q](#) \* *Basis:* Malcolm Baldrige [Q](#) \* *Device:* CTD/Rosette (CTD-RO) [Q](#) \* *Comment:* Section AR21

**3175MB93/38-46** [Q](#) [📄](#) \* *Latitude:* 23.993000 \* *Longitude:* -26.865000 \* *Date/Time:* 1993-08-06T11:24:00 \* *Campaign:* 3175MB93 [Q](#) \* *Basis:* Malcolm Baldrige [Q](#) \* *Device:* CTD/Rosette (CTD-RO) [Q](#) \* *Comment:* Section AR21

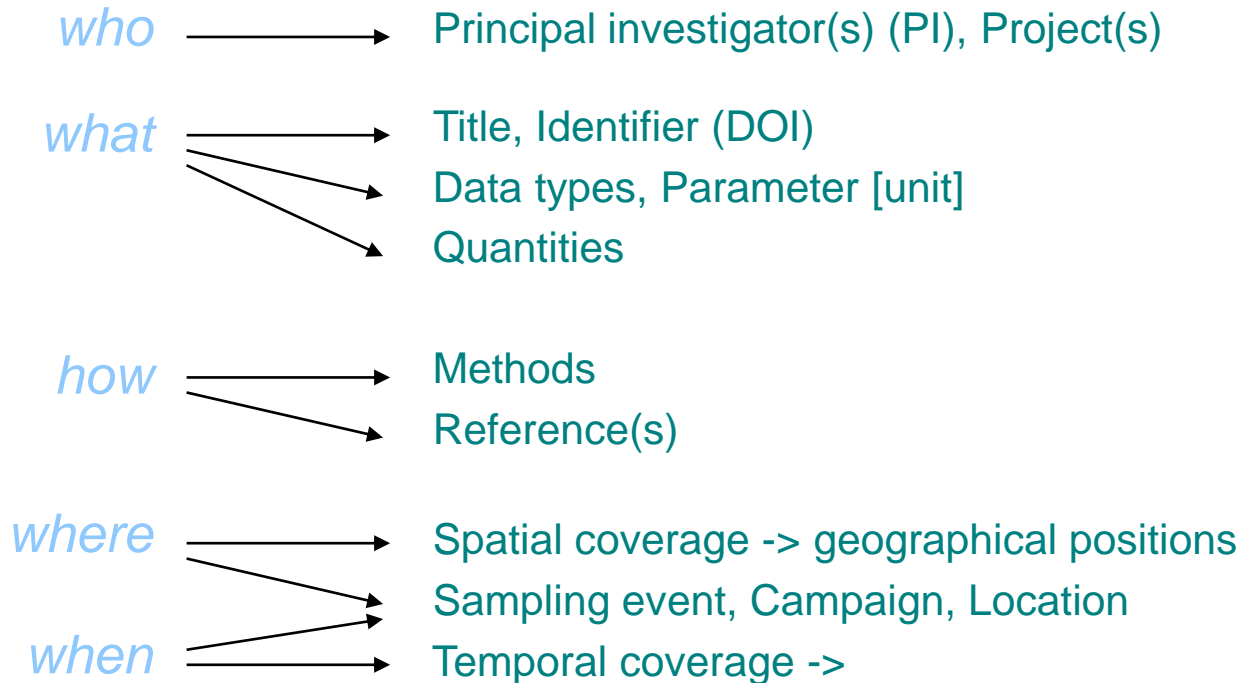
**3175MB93/46-55** [Q](#) [📄](#) \* *Latitude:* 21.007000 \* *Longitude:* -28.465000 \* *Date/Time:* 1993-08-11T09:45:00 \* *Campaign:* 3175MB93 [Q](#) \* *Basis:* Malcolm Baldrige [Q](#) \* *Device:* CTD/Rosette (CTD-RO) [Q](#) \* *Comment:* Section AR21

Parameter(s):

#	Name	Short Name	Unit	Principal Investigator	Method	Comment
1	Event label <a href="#">Q</a>	Event				
2	Date/Time of event <a href="#">Q</a>	Date/Time				
3	Latitude of event <a href="#">Q</a>	Latitude				
4	Longitude of event <a href="#">Q</a>	Longitude				
5	DEPTH, water <a href="#">Q</a>	Depth water	m			Geocode
6	Pressure, water <a href="#">Q</a>	Press	dbar		CTD <a href="#">Q</a>	
7	delta 13C, dissolved inorganic carbon, preindustrial <a href="#">Q</a>	d13C PI	per mil PDB	Olsen, Are <a href="#">Q</a>	Calculated <a href="#">Q</a>	according to Olsen & Ninnemann (2010)
8	Sample, optional label/labor no <a href="#">Q</a>	Lab no				GLODAP station number

License: Creative Commons Attribution 3.0 Unported

# Metadata – describing your data



## What has to be done prior data archival?

---

- › Get the data!
- › Ensure completeness and consistency (reformatting, standardised vocabulary, units)
- › Quality control and quality assurance (scientific expertise)
- › Documentation

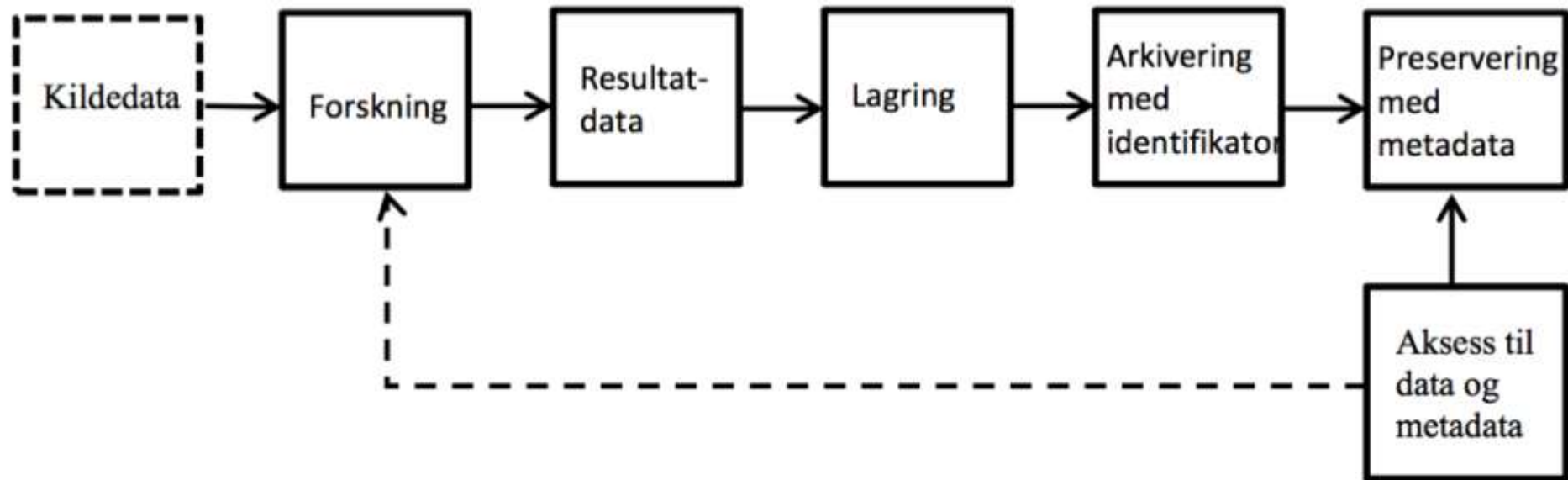
# How is data from UiB's Faculty of Mathematics and Natural Sciences being archived?

---

- › Most of the data is currently not being long-term archived or made available
- › Data is being lost for various reasons
- › 'Geofag evalueringen' criticized data handling procedures
- › Lack of visibility for UiB's scientists and UiB!

# Data management at UiB's Faculty of Mathematics and Natural Sciences?

- › Data management activities within external funded projects (EU, NFR)
- › Data availability varies on the field (->funding)
- › No permanent funded positions for scientific data management
- › GFI / Bjerknes Centre for Climate Research -> Bjerknes Climate Data





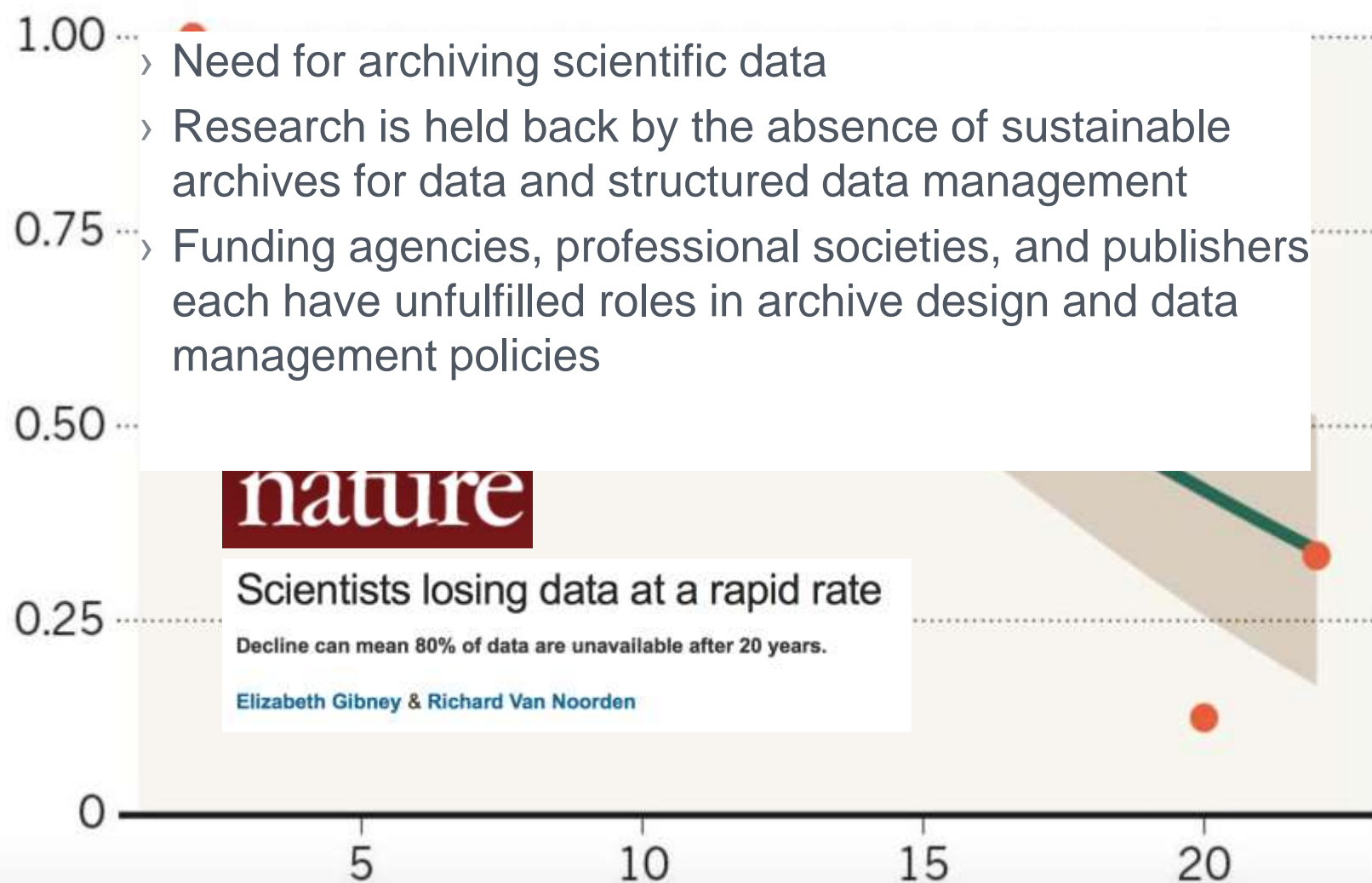


## Not alone...

# MISSING DATA

As research articles age, the odds of their raw data being extant drop dramatically.

Data extant (assuming author responded)



- › Need for archiving scientific data
- › Research is held back by the absence of sustainable archives for data and structured data management
- › Funding agencies, professional societies, and publishers each have unfulfilled roles in archive design and data management policies

**nature**

Scientists losing data at a rapid rate

Decline can mean 80% of data are unavailable after 20 years.

Elizabeth Gibney & Richard Van Noorden



er 2009

inolv



# Good scientific practice in research and scholarship

European Science Foundation (ESF), 2000

## Data accumulation, handling, and storage

---

36. Data are produced at all stages in experimental research and in scholarship. Data sets are an important resource, which enable later verification of scientific interpretations and conclusions. They may also be the starting point for further studies. It is vital, therefore, that all primary and secondary data are stored in a secure and accessible form.

37. Institutions may pay particular attention to documenting and archiving original research and scholarship data. Several codes of good practice recommend a minimum period of 10 years, longer in the case of especially significant or sensitive data. National or regional discipline-based archives should be considered where there are practical or other problems in storing data at the institution where the research was conducted.

# Principles for dissemination of scientific data (ISCU/CODATA, 2000)

**4. Scientific advances rely on full and open access to data.** Both science and the public are well served by a system of scholarly research and communication with minimal constraints on the availability of data for further analysis. The tradition of full and open access to data has led to breakthroughs in scientific understanding, as well as to later economic and public policy benefits. The idea that an individual or organization can control access to or claim ownership of the facts of nature is foreign to science.

**5. *The interests of database owners must be balanced with society's need for open exchange of ideas.*** Given the substantial investment in data collection and its importance to society, it is equally important that data are used to the maximum extent possible. Data that were collected for a variety of purposes may be useful to science. Legal foundations and societal attitudes should foster a balance between individual rights to data and the public good of shared data.



## Norway (+30 OECD lands) signed the Declaration on Access to Research Data from Public Funding, January 2004

Source: <http://acts.oecd.org>



### Declaration on Access to Research Data from Public Funding

30 January 2004 - C(2004)31/REV1



THE GOVERNMENTS<sup>[1]</sup> of Australia, Austria, Belgium, Canada, China, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Russian Federation, the Slovak Republic, the Republic of South Africa, Spain, Sweden, Switzerland, Turkey, the United Kingdom, and the United States:

RECOGNISING that an optimum international exchange of data, information and knowledge contributes decisively to the advancement of scientific research and innovation;

RECOGNISING that open access to, and unrestricted use of, data promotes scientific progress and facilitates the training of researchers;

RECOGNISING that open access will maximise the value derived from public investments in data collection efforts;

RECOGNISING that the substantial increase in computing capacity enables vast quantities of digital research data from public funding to be put to use for multiple research purposes by many research institutes of the global science system, thereby substantially increasing the scope and scale of research;

RECOGNISING the substantial benefits that science, the economy and society at large could gain from the opportunities that expanded use of digital data resources have to offer, and recognising the risk that undue restrictions on access to and use of research data from public funding could diminish the quality and efficiency of scientific research and innovation;

RECOGNISING that optimum availability of research data from public funding for developing countries will enhance their participation in the global science system, thereby contributing to their social and economic development;

RECOGNISING that the disclosure of research data from public funding may be constrained by domestic law on national security, the protection of privacy of citizens and the protection of intellectual property rights and trade secrets that may require additional safeguards;

RECOGNISING that on some of the aspects of the accessibility of research data from public funding, additional measures have been taken or will be introduced in OECD countries and that disparities in national regulations could hamper the optimum use of publicly funded data on the national and international scales;

CONSIDERING the beneficial impact of the establishment of OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (1980, 1985 and 1990) and the OECD Guidelines for the Security of Information Systems and Networks (1992, 1997 and 2002) on international policies for access to digital data;

#### DECLARE THEIR COMMITMENT TO:

Work towards the establishment of access regimes for digital research data from public funding in accordance with the following objectives and principles:

**Openness:** balancing the interests of open access to data to increase the quality and efficiency of research and innovation with the need for restriction of access in some instances to protect social, scientific and economic interests.

**Transparency:** making information on data-producing organisations, documentation on the data they produce and specifications of conditions attached to the use of these data, available and accessible internationally.

**Legal conformity:** paying due attention, in the design of access regimes for digital research data, to national legal requirements concerning national security, privacy and trade secrets.

**Formal responsibility:** promoting explicit, formal institutional rules on the responsibilities of the various parties involved in data-related activities pertaining to authorship, producer credits, ownership, usage restrictions, financial arrangements, ethical rules, licensing terms, and liability.

## OECD Principles and Guidelines for Access to Research Data from Public Funding (2007)

The rapid development in computing technology and the Internet have opened up new applications for the basic *sources* of research — the base material of *research data* — which has given a major impetus to scientific work in recent years.

Databases are rapidly becoming an essential part of the infrastructure of the global science system. The international Human Genome Project is but one good example of a large-scale endeavour in which openly accessible information is being used successfully by many different users, all over the world, for a great variety of purposes.



# OECD Principles and Guidelines for Access to Research Data from Public Funding (2007)

- Promote a culture of openness and sharing of research data among the public research communities within member countries and beyond;
- Stimulate the exchange of good practices in data access and sharing;
- Raise awareness about the potential costs and benefits of restrictions and limitations on access to and the sharing of research data from public funding;
- Highlight the need to consider data access and sharing regulations and practices in the formation of member countries' science policies and programmes;

The *Principles and Guidelines* should assist governments, research support and funding organisations, research institutions and researchers themselves in dealing with the barriers and challenges in improving the international sharing of, and access to, research data.

# Anbefalinger fra policyen

## Forskningsdata bør

- arkiveres på en sikker måte
- gjøres tilgjengelige for alle relevante brukere, under like vilkår
- gjøres tilgjengelig så tidlig som mulig, men en embargoperiode for å publisere er ok
- utstyres med metadata basert på internasjonale standarder
- utstyres med lisenser som legger så få begrensninger som mulig på tilgang, gjenbruk og videredistribusjon
- gjøres tilgjengelig til lavest mulig kostnad; maksimalt de faktiske kostnadene knyttet til tilgjengeliggjøring
- utstyres med en langtidsplan



Photo: Shutterstock/Marynchenko Oleksandr

```
<!DOCTYPE html PUBLIC "-//W3C//  
<html xmlns="http://www.w3.org/  
<head>  
<meta http-equiv="Content-  
<meta http-equiv="Content-  
<meta http-equiv="Content-  
<title>Document Title</titl  
<link rev="made" href="mail  
<link rev="start" href="./'  
<style type="text/css" medi
```

Photo: Shutterstock/Alexey V Smirnov



Illustration: Creative Commons

## Guidelines on the Implementation of Open Access to Scientific Publications and Research Data in Projects supported by the European Research Council under Horizon 2020

---

- › Open access to publications and data
- › ERC grants must ensure open access to all peer-reviewed scientific publications relating to its results.
- › Open access to research data
  - Open Research Data Pilot (ORD Pilot) – mandatory!
  - Beneficiaries must deposit the data in a research data repository so that it is **possible to access, mine, exploit, reproduce and disseminate the data**
  - **Provide a Data Management Plan (DMP)**



## Publishers

---

Data underlying scientific publications have to be open!  
Being enforced by rejecting papers

## Reality check on reproducibility



**A survey of *Nature* readers revealed a high level of concern about the problem of irreproducible results. Researchers, funders and journals need to work together to make research more reliable.**

25 May 2016      **2/3 of all publications can not be reproduced!**

# Is just archiving data enough?

---



### Guidelines on FAIR Data Management in Horizon 2020

Aim to make research data **FAIR**

- **Findable**
- **Accessible**
- **Interoperable**
- **Reusable**
- Good research data management is not a goal in itself, but rather the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse.
- "as open as possible, as closed as necessary" and focuses on encouraging sound data management as an essential part of research best practice.

Source: <https://data.europa.eu>

### Guidelines on FAIR Data Management in Horizon 2020

---

#### Data management plan (DMP)

As part of making research data findable, accessible, interoperable and re-usable (FAIR), a DMP should include information on:

- the handling of research data during and after the end of the project
- what data will be collected, processed and/or generated
- which methodology and standards will be applied
- whether data will be shared/made open access and
- how data will be curated and preserved (including after the end of the project).



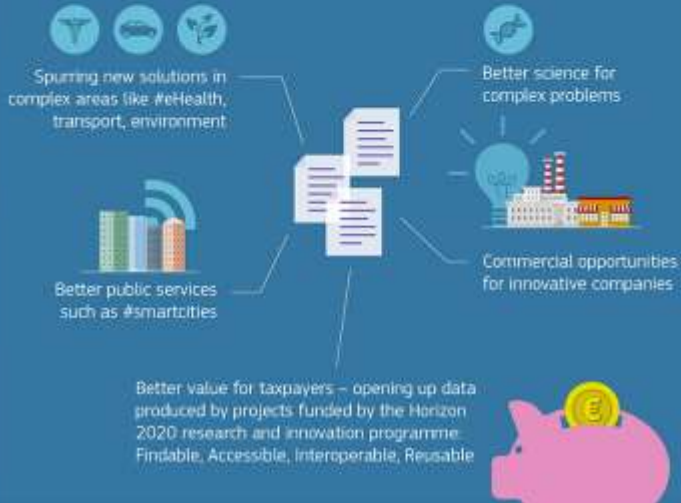
#DigitalSingleMarket



## EUROPEAN CLOUD INITIATIVE

UNLOCKING THE POWER OF BIG DATA FOR OPEN SCIENCE

### OPPORTUNITIES



## EUROPEAN OPEN SCIENCE CLOUD

BRINGING TOGETHER CURRENT AND FUTURE DATA INFRASTRUCTURES



## EUROPEAN DATA INFRASTRUCTURE

UNLOCKING THE VALUE OF BIG DATA: DIGITAL BY DEFAULT





# How to archive data: Technical data organisation

## File systems

*disadvantage: low consistency of data*

*advantage: fast & cheap archiving procedure (on a short run)*

## Relational databases (RDBs)

*disadvantage: work intensive archiving procedure,  
needs high degree of data organization  
usage for mass data is limited*

*advantage: high consistency of data,  
low costs for data curation,  
good retrieval qualities*

## Mixed

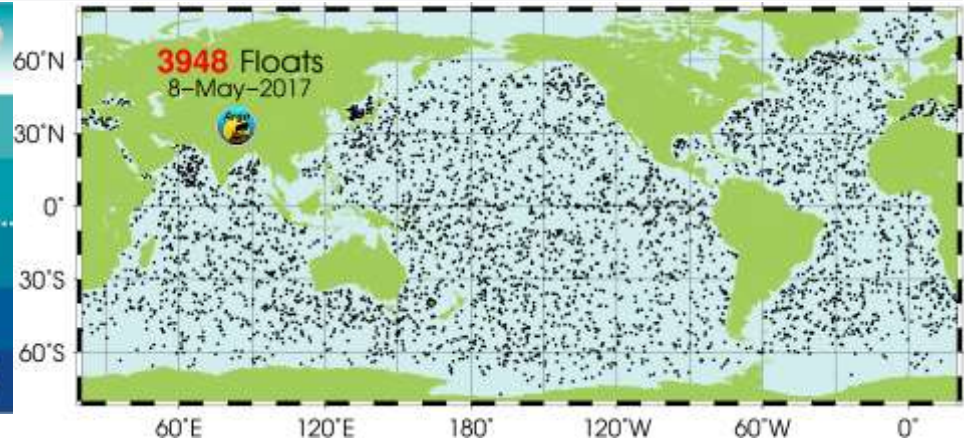
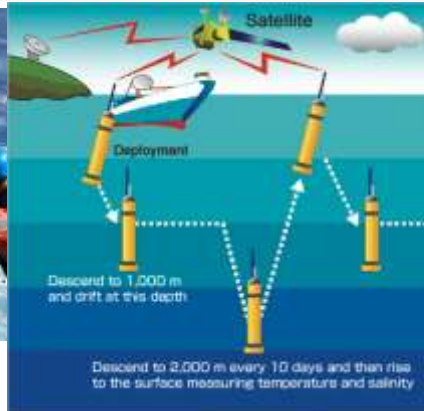
*Relational database -> geocoded data & metadata*

*File system -> mass data (geophysical data, pictures, films)*

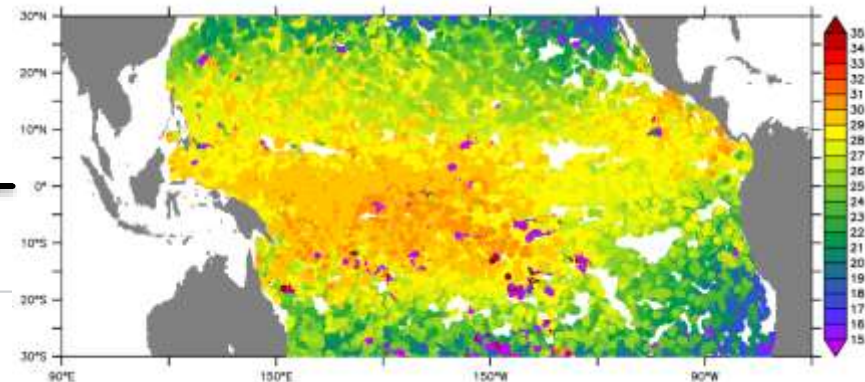
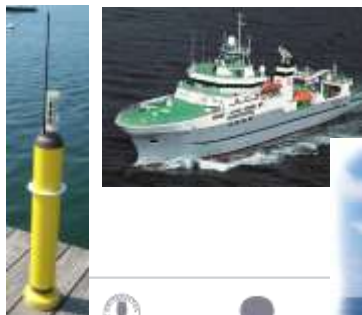




# FAIR: Integrating Data and Information Across Observing Systems -> Fitness for Use



Data availability via DAP (Data Access Protocol) services e.g. OPeNDAP allowing interoperable access to data from various observing networks





# Interested in becoming a data scientist? According to Harvard Business Review:

DATA

# Data Scientist: The Sexiest Job of the 21st Century

by **Thomas H. Davenport** and **D.J. Patil**

FROM THE OCTOBER 2012 ISSUE

 SUMMARY

 SAVE

 SHARE

 COMMENT 5

 TEXT SIZE

 PRINT

 **\$8.95**  
BUY COPIES

**W**hen Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't seeking out connections with the people who were already on the site at

## WHAT TO READ NEXT

[Big Data: The Management Revolution](#)

[5 Essential Principles for Understanding Analytics](#)

[Data Scientists Don't Scale](#)

VIEW MORE FROM THE  
October 2012 Issue





# Formal requirements by

---

- › Funding agencies
- › Publishers
- › Researchers
- › Society

# Hvordan vil vi følge opp policyen?

- Kostnader i prosjektene kan inkluderes i søknader om støtte fra Forskningsrådet
  - Kostnader ved å gi åpen tilgang til resultatdata
  - Kostnader ved å få tilgang til kilde-data
- Utreder krav om en datahåndteringsplan ved kontraktsinngåelse for prosjekter som får støtte
- Oppfordre forskningsinstitusjonene til, i samarbeid med egne forskere, å lage retningslinjer for hvilke data som skal lagres og tilgjengeliggjøres og tydeliggjøre hvilke løsninger de anbefaler at forskerne benytter for håndtering av ulike datasett.
- Støtte infrastrukturer som bidrar til
  - åpen tilgang til offentlig finansierte forskningsdata
  - sikring av viktige norske dataserier



Photo: Shutterstock



Photo: Digital Vision

## H2020 Programme

### Guidelines on FAIR Data Management in Horizon 2020

---

Costs related to open access to research data in Horizon 2020 are eligible for reimbursement during the duration of the project under the conditions defined in the H2020 Grant Agreement  
Provides template for a DMP (similar to the British Data Curation Centre)

Source: <https://data.europa.eu>

## Structured data management at GFI

---

Data management is a resource for funding:

- › Funding agencies favour projects where DM is allocated with a financial post (re-assurance)
  - Example UiB/GFI since 2011 data management activities
    - 20+ funded projects
    - 28+ mill NOK
    - 5 (soon 6) data managers on soft money

## Structured data management at GFI

---

Data management is a resource for visibility:

- Bjerknes Climate Data Centre (2014)
- Hosting data management office for RI ICOS OTC (2016)
- Data services for data products (SOCAT, GLODAP)
- IOC UNESCO IODE Global Data Assembly Centre (2017?)
- Committee work => 14 since 2010 (IOC UNESCO (IOCCP, GOSUD), OECD (GSF), RIs, global initiatives (GOA-ON, SOOS), data archives, project level)
- Many papers (20+) and acknowledgements



REMOTE SENSING

PALEO

MODEL OUTPUT

OCEANOGRAPHY

ATMOSPHERE

DATA PRODUCTS

NEWS

Currently there are no events.

>> About

## WELCOME TO THE BJERKNES CLIMATE DATA CENTRE

THE BJERKNES CLIMATE DATA CENTRE (BCDC) HAS BEEN LAUNCHED IN JANUARY 2015 AND SOME PARTS OF THE WEBSITE ARE STILL UNDER CONSTRUCTION

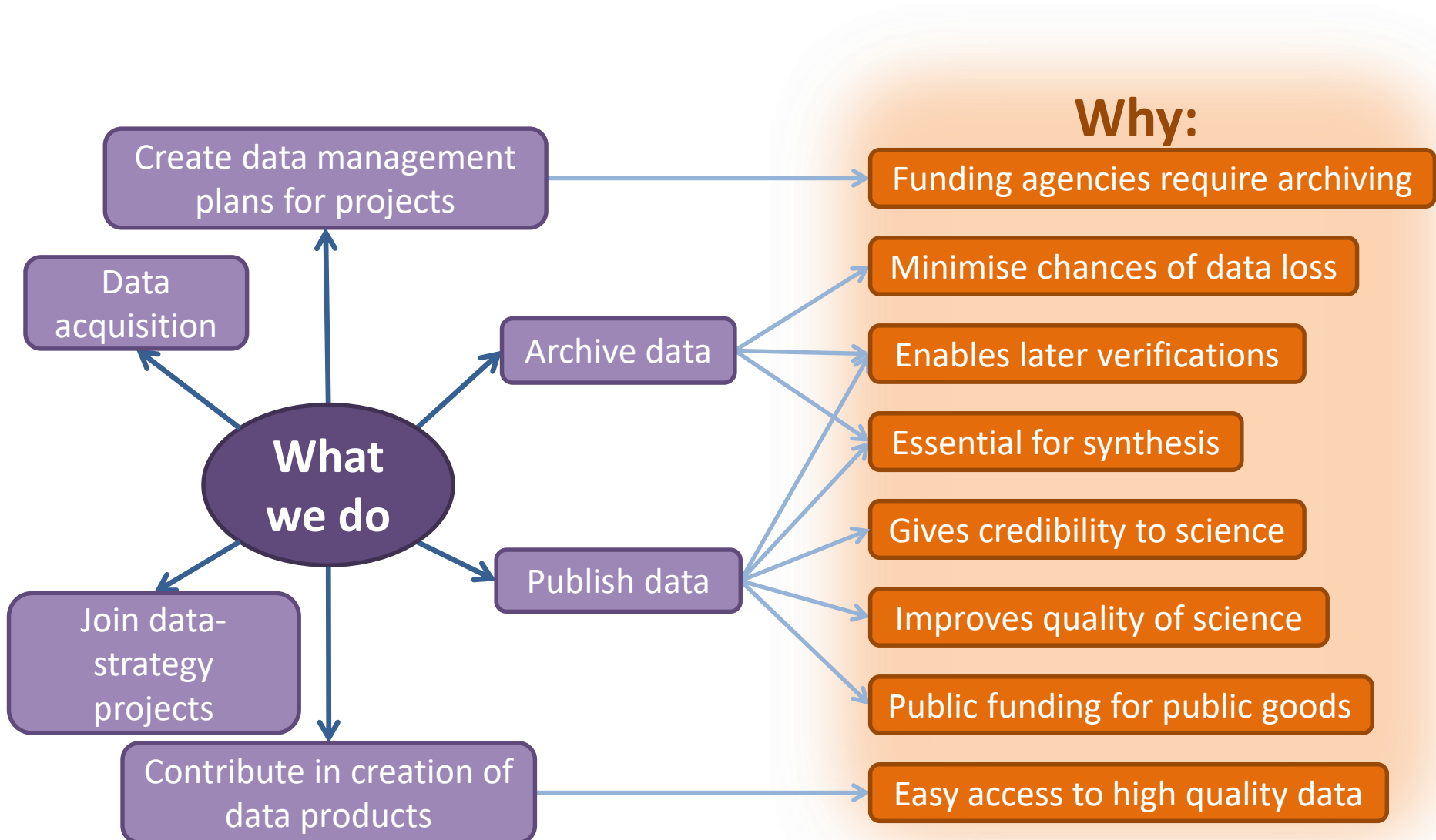
### ABOUT

The BCDC has the following aims:

1. To serve the data obtained, and the data syntheses assembled, by researchers within the Bjerknes Centre for Climate Research cluster. All data from the different disciplines (e.g. geology, oceanography, biology, model community) will be archived, interconnected and made publicly available by the BCDC. It will, however, be open for all interested scientists independent of institution.
2. To provide an online access portal to climate data of all kinds. The data center will be embedded in the worldwide data access network established by the International Council for Science called World Data System (ICSU WDS, former ICSU World Data Center System, <http://www.icsu-wds.org/>) and follow the idea of having a common globally interoperable distributed data system within the climate (change) community. We will also cooperate with data



# What do we do?



# Data Sources

Harvesting Software for Metadata

Data catalogues

Bjerknes Climate Data Centre



Data request links back to the respective source

Credit goes to the data source!

Model output

Remote Sensing

Physical Oceanography

Chemical Oceanography

Marine and Limnic Geology

NRT Oceanography



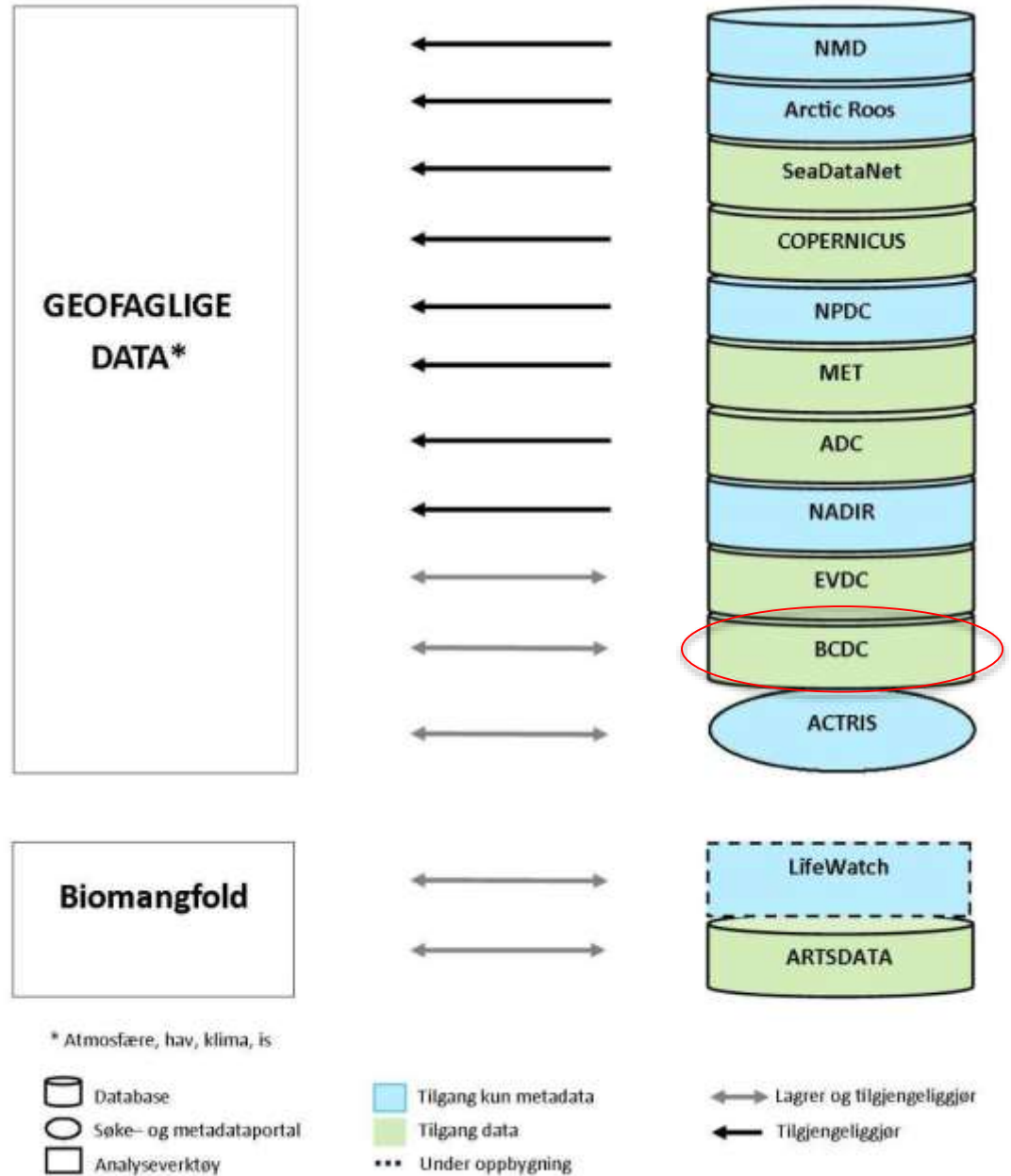
## How does it (technically) work?

ICSU WDS member in 2017

**Main aim: do not re-invent the wheel – focus on services**

- › use existing infrastructure and adjust them to our needs
- › embedded in the worldwide data access network established by the International Council for Science called World Data System
- › data search provides metadata
- › outsourced infrastructure through collaboration with PANGAEA
- › collaboration with other data centres that are not part of ICSU WDS and make their metadata available (e.g. NORSTORE, NMD)
- › streamline our data to other metadata inventories





Figur 4: Kategorisering av tjenestene oppsummert i tabell 9 (naturvitenskap), i "database", "søke- og metadataportal" eller "analyseverktøy", samt informasjon om hvorvidt tjenestene gir tilgang til data (og metadata) eller kun metadata.

## Vision for structured data management at UiB

---

- › Clear need for a structured and sustainable framework **integrated** in the field of natural sciences (currently not performed by IT department or UiB library)
- › Should become part of the basic infrastructure and services should be available to scientists on a long-term basis (not 100% dependence on soft money)
- › Need for a UiB data policy!
- › Need for professional data managers addressing various fields of natural sciences
- › No need for own long-term archive!

## How shall it be financed?

---

- Eligible costs (external funding = soft money)
- In-kind contribution
- Overheads

## What we will gain?

---

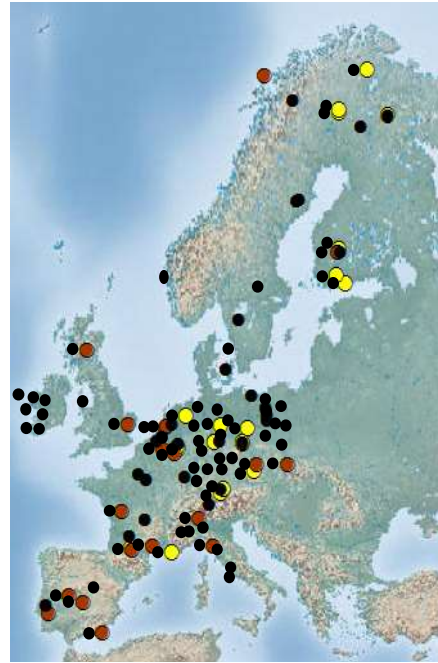
- Visibility (science and infrastructure)
  - Credibility
  - Support for scientists on several levels
  - Fulfil legal and societal obligations
  - Expertise
- > Become attractive collaborators and attract funding for scientists

# The potential ICOS station RI-network

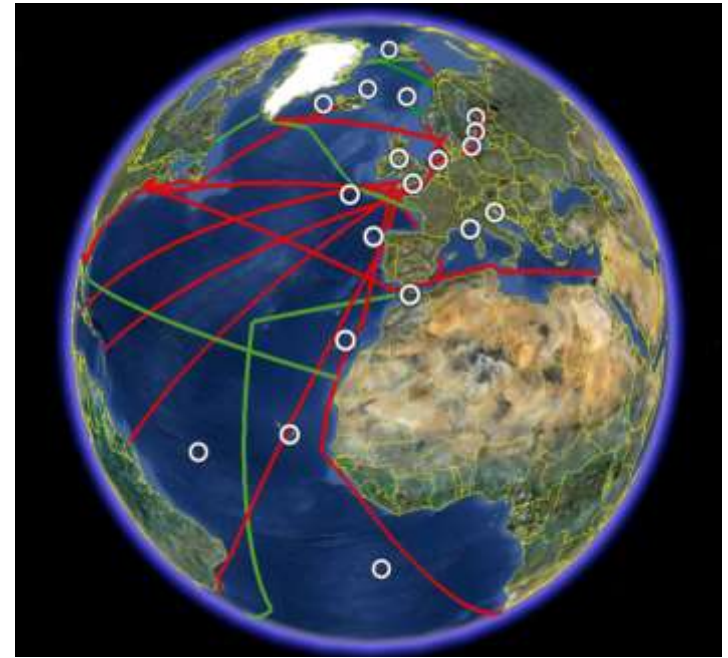
Atmosphere



Ecosystems



Oceans



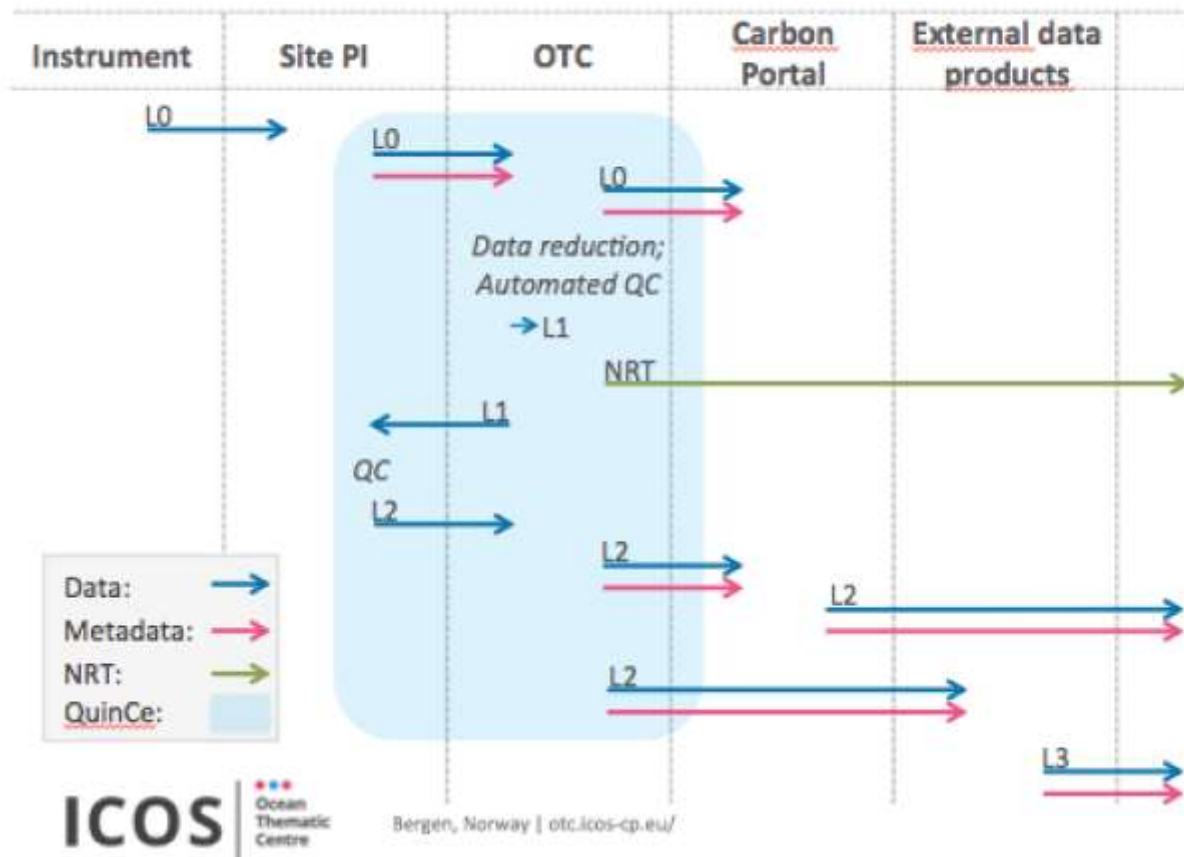
ICOS Norway and OTC:  
Funded by NFR approx. 93 mill NOK (2016-2021)

Coordinator: Truls Johannessen (UiB)  
Project management: Abdir Omar (Uni Research)  
Data management: Benjamin Pfeil

# ICOS OTC Data management

Backbone of ICOS (>50 % of the OTC deliverables)

## Data Lifecycle within OTC



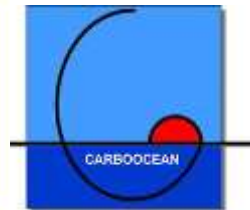
deliverables ensure:  
data exchange with

partners e.g. British  
Model

products including QC,

# History of marine carbon data management at BCCR/UiB

- Dedicated funding for marine carbon research and data management through the EU projects CARBOOCEAN (2005-2009) and CARBOCHANGE (2011-2015) coordinated by Christoph Heinze





# Community efforts with strong European contribution and international support



## A multi-decade record of high quality fCO<sub>2</sub> data in version 3 of the Surface Ocean CO<sub>2</sub> Atlas (SOCAT)

D. C. E. Bakker<sup>1</sup>, B. Pfeil<sup>2,3</sup>, C. S. Landa<sup>2,3</sup>, N. Metzl<sup>4</sup>, K. M. O'Brien<sup>5,6</sup>, A. Olsen<sup>2,3</sup>, K. Smith<sup>5,6</sup>, C. Cosca<sup>5</sup>, S. Harasawa<sup>7</sup>, S. D. Jones<sup>8</sup>, S. Nakaoka<sup>7</sup>, Y. Nojiri<sup>7</sup>, U. Schuster<sup>8</sup>, T. Steinhoff<sup>9</sup>, C. Sweeney<sup>10</sup>, T. Takahashi<sup>11</sup>, B. Tilbrook<sup>12,13</sup>, C. Wada<sup>7</sup>, R. Wanninkhof<sup>14</sup>, S. R. Alin<sup>5</sup>, L. Barbero<sup>15</sup>, N. R. Bates<sup>16</sup>, A. A. Bianchi<sup>17</sup>, F. Bonou<sup>18</sup>, J. Boutin<sup>4</sup>, Y. Bozec<sup>19,20</sup>, W.-J. Cai<sup>21</sup>, R. D. Castle<sup>14</sup>, L. Chen<sup>22,23</sup>, M. Chierici<sup>24,25</sup>, K. Currie<sup>26</sup>, W. Evans<sup>5,27,28</sup>, C. Featherstone<sup>14</sup>, R. A. Feely<sup>5</sup>, A. Fransson<sup>29</sup>, N. Greenwood<sup>30</sup>, L. Gregor<sup>31</sup>, S. Hankin<sup>5</sup>, N. J. Hardman-Mountford<sup>32</sup>, J. Harlay<sup>33</sup>, J. Hauck<sup>34</sup>, M. Hoppema<sup>34</sup>, M. Humphreys<sup>35</sup>, C. W. Hunt<sup>36</sup>, B. Huss<sup>14</sup>, J. S. P. Ibánhez<sup>37,18</sup>, T. Johannessen<sup>2,3,38</sup>, R. Keeling<sup>39</sup>, V. Kitidis<sup>40</sup>, A. Körtzinger<sup>9</sup>, A. Kozyr<sup>41</sup>, E. Krasakopolou<sup>42</sup>, A. Kuwata<sup>43</sup>, P. Landschützer<sup>44</sup>, S. K. Lauvset<sup>2,3</sup>, N. Lefèvre<sup>45,46</sup>, C. LoMonaco<sup>4</sup>, A. B. Manke<sup>5</sup>, J. T. Mathis<sup>5</sup>, L. Merlivat<sup>4</sup>, Millero, F. J.<sup>47</sup>, P. Monteiro<sup>31</sup>, D. Munro<sup>10</sup>, A. Murata<sup>48</sup>, T. Newberger<sup>10</sup>, A. M. Omar<sup>38,3,2</sup>, T. Ono<sup>49</sup>, K. Paterson<sup>12</sup>, D. Pierrot<sup>15</sup>, L. L. Robbins<sup>50</sup>, C.L. Sabine<sup>5</sup>, S. Saito<sup>51</sup>, J. Salisbury<sup>36</sup>, R. Schlitzer<sup>34</sup>, B. Schneider<sup>52</sup>,

UiB had the competence in science, project management and data management which made us an attractive partner to host the RI ICOS OTC!

## What do I want to achieve?

---

1. Raising the awareness of the fact that observational data obtained at UiB are currently not being handled on a systematic bases
2. Publish data and initiate data rescue activities → scientific heritage
3. Common data policy for UiB where PIs are obliged to make their data availability
4. Sustainable funding by consistently joining all relevant proposals
5. Make it sustainable by having professional data managers that will archive data using existing infrastructure



*'The value of data lies in their use.'*

Bits of Power, Issues on Global Access to Scientific Data, US National  
Research Council 1997

*Thank you!*

<https://drive.google.com/file/d/0B1dyZ3Lmzn8SNnUybXZOTERaUTA/view>