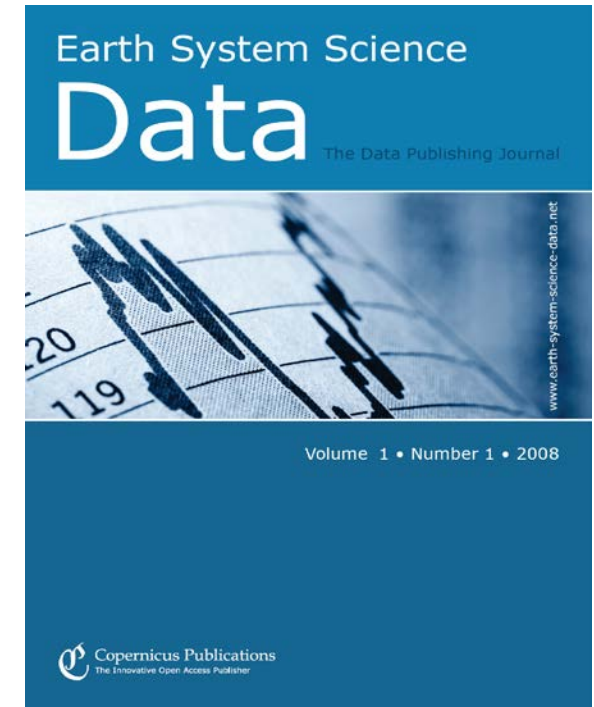


Earth Systems Science Data – a Journal and its Ecosystem

Hans Pfeiffenberger

Alfred-Wegener-Institut, Helmholtz Zentrum für Polar- and Meeres-Forschung,

DINI Jahrestagung 2017, Göttingen, 2017-10-05



Publishing - with a capital P

- to **publish**: make publicly available, e.g. upload to website
- to **Publish** - as in Scientific Publishing :
 - is a quite **formal, “ritualised” process**
requires systematic **3rd party scrutiny**
=> QA supposedly yields **higher quality**
 - establishes **priority**
 - quality + priority => **reputation => willingness to share**
 - published items become part of **“The Records of Science”**

(Data) Publishing - with a capital P

- **Standardized, well known process**
=> TRUST => we can build on the work of others !!
- **Apply these mechanisms to data**
 - **get all the benefits !! ??**
 - **thus: “Data Publishing with a capital P”**
- **In a broader sense, this is about:**
To make data a legitimate part of research culture

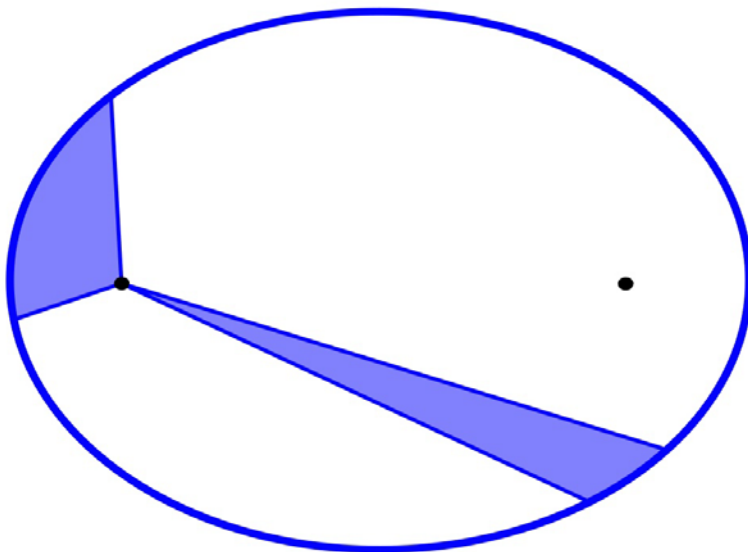
Agenda

- **A little bit of “historical perspective”**
- **ESSD, a Data Publishing journal, est. 2008**
with a note on presentation of data
- **Bonusmaterial? Linked Scientific Information**
(all Capital Letters) – but what are the objects?

An impression from history of science

Modern Science is based on data
– since Renaissance!

- **1606 - 1618: Kepler's Laws**
 - from Tycho Brahe's quality data
- **1684 – 1687 Newton De Motu – Principia**
 - **explained (!) Kepler's laws**



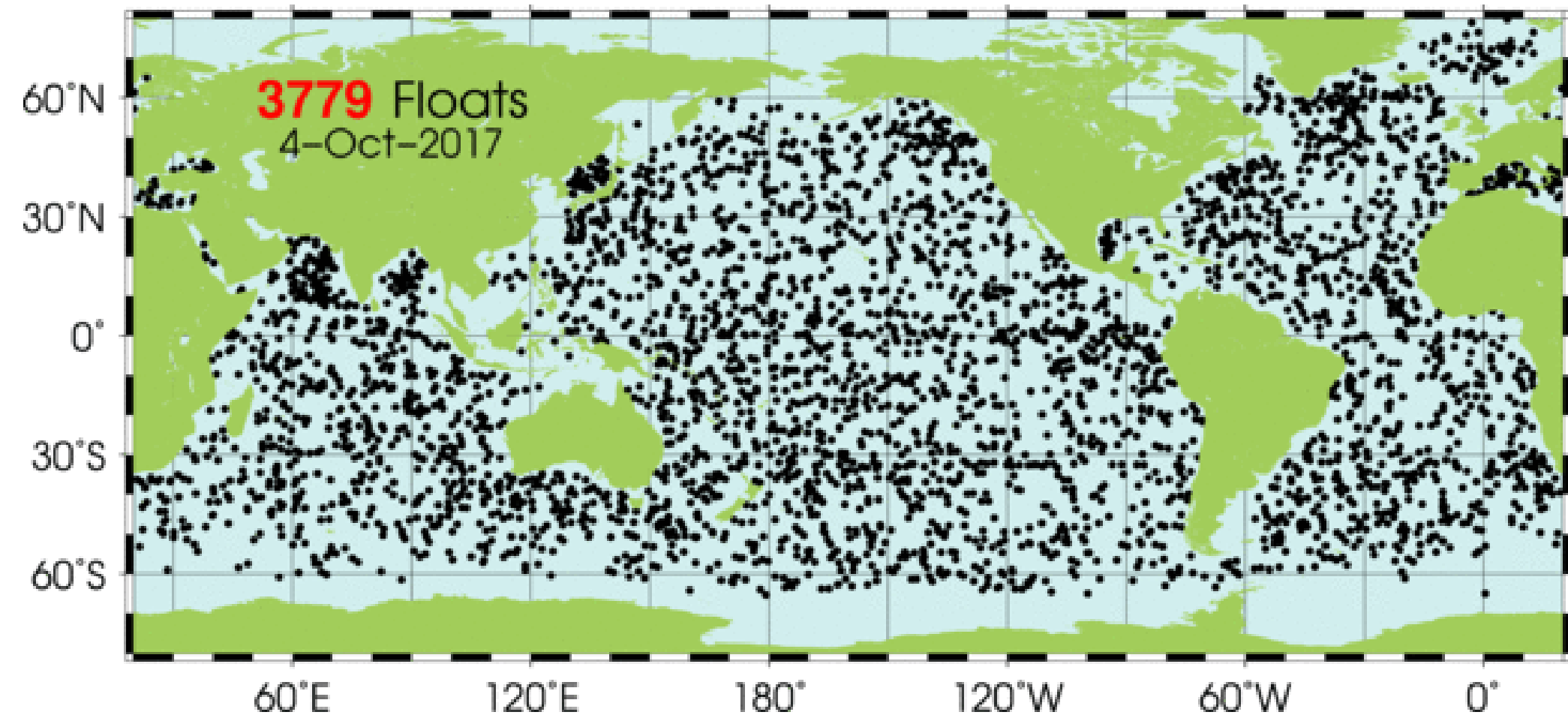
$$F = ma$$

$$F \sim \frac{mM}{r^2}$$

Planet	T	d	T^2	d^3	T^2/d^3
Merkur	0,241	0,387	0,058081	0,057960603	1,002077221
Venus	0,615	0,723	0,378225	0,377933067	1,000772446
Erde	1	1	1	1	1
Mars	1,881	1,524	3,538161	3,539605824	0,999591812
Jupiter	11,863	5,203	140,730769	140,8515004	0,999142846
Saturn	29,458	9,555	867,773764	872,3526289	0,994751131

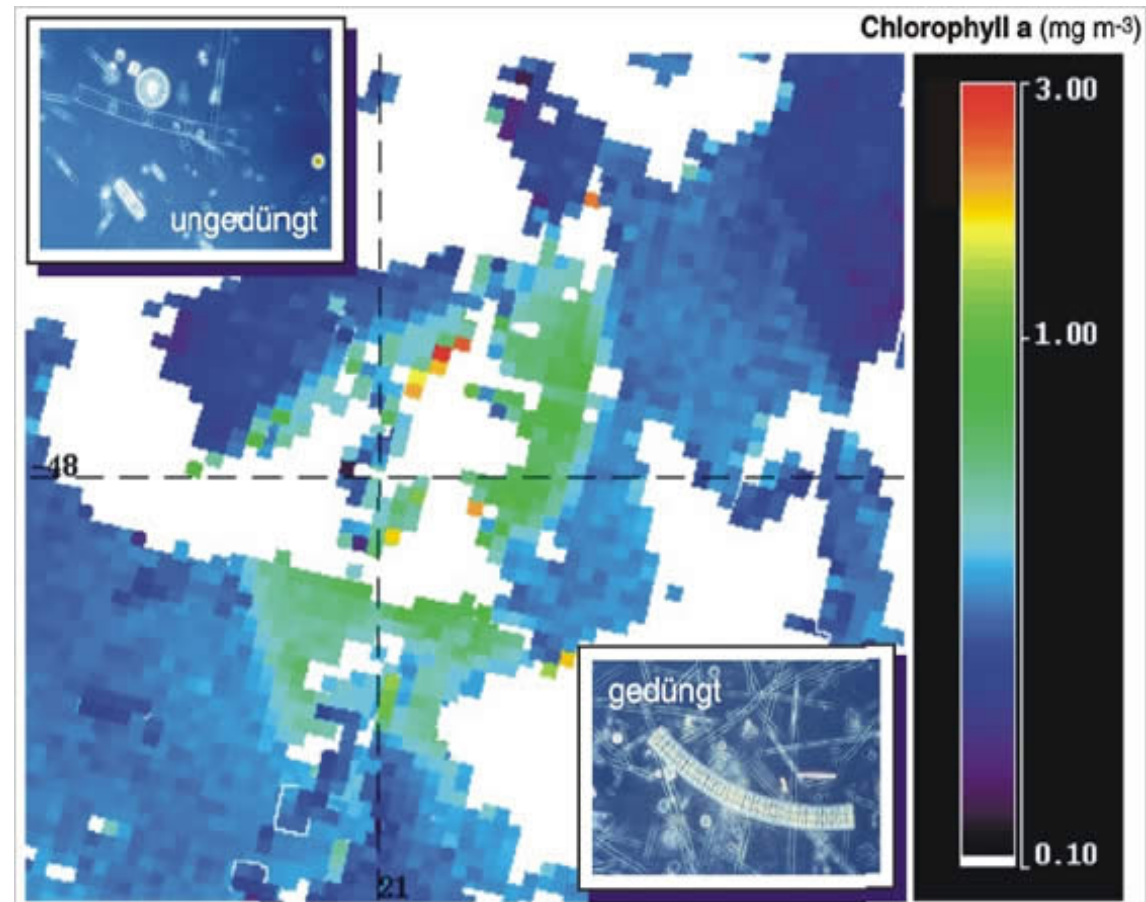
T = siderische Umlaufzeit in trop. Jahren d = große Halbachse in astronomischen Einheiten (Abstand Erde–Sonne)

The biggest experiment on this planet - ARGO



An important, “typical” Experiment

- EISENEX / EIFEX : Two expeditions of “**Polarstern**” :
With a few tons of iron fertilizer, south of Capetown
- EIFEX (2004):
 - 54 scientists and students from
 - 14 institutes and 3 companies from
 - 7 EU countries and South Africa
 - Oceanographers
 - Biologists
 - Chemists.....
- “**Biogeochemistry**”
- **+ Satellite observations !**



ESSD – Data Publishing in practise

2013: CO above Troll Station, Original Data

BAS microwave radiometer CO profiles acquired at Troll station, Antarctica between Feb 2008 and Jan 2010
 Contact: Patrick Espy, tel: +47 73 55 10 95, email: patrick.espy@ntnu.no

 date [UT]: 2009-10-19 10:44:06
 apriori contribution: The profile is most reliable where the contribution from the a priori profile is less than approx. Negative values are a scaling artifact and should be regarded as close to 0.

The 2-sigma systematic errors provided have been determined using perturbation calculations:

temperature error: error induced by the temperature profile (estimated error = 5K) needed as additional information for the retrieval, mainly random
 calibration error: error induced by the calibration of the measured spectrum (estimated error = 10 percent), can be systematic
 spectroscopy error: we used lineintensity from HITRAN 2004 with an estimated error of 2 percent, systematic
 channel shape error: uncertainty due to the use of a modified channel response function in the retrieval in order to correct for an instability in one of the radiometers local oscillators after 2008-08-09, systematic
 Error from measurement noise [K]: 0.1510, random
 Smoothing error: This error only needs to be considered if the profiles of the BAS radiometer are compared to profiles with a significantly larger vertical resolution. For such a comparison the better way would be to convolve the high-resolution profile with the AVK of the retrievals.

Sum of errors: To build the sum of certain errors they are added up as follows $\sqrt{\text{error1}^2 + \text{error2}^2}$

pressure [hPa]	altitude [km]	vmr [ppmv]	apriori contribution [percent]	temperature error [ppmv]	calibration error [ppmv]	spectroscopy error [ppmv]
0.749894	50.679	0.060	-5.939	0.003	0.048	0.234
0.562341	53.021	0.065	-20.151	0.002	0.056	0.319
0.421697	55.337	0.072	-27.600	0.002	0.061	0.349
0.316228	57.609	0.080	-29.442	0.004	0.067	0.298

Sun-earth Interactions measurements carried out in order to study the dynamical context. [Storage](#)
 > The data set covers the period from February 2008 to January 2010, however, due to very low CO concentrations [Constraints](#)

- General Information
- Submission
- Review

Abstract. This paper presents mesospheric carbon monoxide (CO) data acquired by the ground-based microwave radiometer of the British Antarctic Survey (BAS radiometer) stationed at Troll station in Antarctica (72° S, 2.5° E, 1270 a.m.s.l.). The data set covers the period from February 2008 to January 2010, however, due to very low CO

ESSD Principles and Criteria

- ESSD expects data to be at a repository and be
 - **Open Access, static, with a DOI**
- ESSD expects **authors** to describe **in the article**
 - **provenance, methods, limitations, estimates of error**
- ESSD expects **reviewers** to
 - **actually look at the data**
 - **assess consistency** of article and dataset

Fluxes of sedimenting material from sediment traps in the Atlantic Ocean

S. Torres-Valdés¹, S. C. Painter¹, A. P. Martin¹,
R. Sanders¹, and J. Felden²

¹Ocean Biogeochemistry and Ecosystems Research Group,
Southampton, SO14 3ZH, UK

²Center for Marine Environmental Sciences, Universität Bremen,
Bremen, Germany

Review Status

This discussion paper is under review for the journal Earth System Science Data (ESSD).

A huge work to find, assess, collate (quality) data;

24 out of 43 text pages are source data references!

Abstract. We provide a data set assemblage of directly observed and derived fluxes of sedimenting material (total mass, POC, PON, BSiO₂, CaCO₃, PIC and lithogenic/terrigenous fluxes) obtained using sediment traps. This data assemblage contains over 5900 data points distributed across the Atlantic, from the Arctic Ocean to the Southern Ocean. Data from the Mediterranean Sea are also included. Data were compiled from a variety of sources: data repositories (e.g., BCO-DMO, PANGAEA), time series sites (e.g., BATS, CARIACO), published scientific papers and data provided by originating PI's. All sources are specified within the combined data set. Data from the World Ocean Atlas 2009 were extracted to coincide with flux

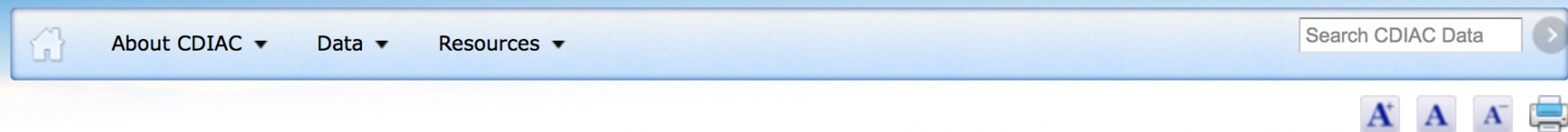
2012, 13, 14, 15, 2016: Nature Climate Change, ESSD and CDIAC

	A	B	C	D	E	F	G	
1	Terrestrial CO₂ sink (positive values represent a flux from the atmosphere to the land)							
2	All values in petagrams of carbon per year (PgC/yr), for the globe. For values in carbon dioxide (CO ₂), multi							
3	1PgC = 1 petagram of carbon = 1 billion tonnes C = 1 gigatonne C = 3.67 billion tonnes of CO ₂							
4	Cite as:							
5	CLM4CN	Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence,						
6	HYLAND	Levy, P. E., M. G. R. Cannell, et al. (2004). "Modelling the impact of future changes in clim						
7	LPJ-GUESS	Smith, B., I. C. Prentice, et al. (2001). "Representation of vegetation dynamics in the mode						
8	LPJ	Sitch, S., B. Smith, et al. (2003). "Evaluation of ecosystem dynamics, plant geography and						
9	O-CN	Zaehle, S., P. Ciais, et al. (2011). "Carbon benefits of anthropogenic reactive nitrogen offs						
10	ORCHIDEE	Krinner, G., N. Viovy, et al. (2005). "A dynamic global vegetation model for studies of the						
11	SDGVM	Woodward, F. I. and M. R. Lomas (2004). "Vegetation dynamics - simulating responses to						
12	JULES	Clark, D. B., L. M. Mercado, et al. (2011). "The Joint UK Land Environment Simulator (JULE						
13	VEGAS	Zeng, N., A. Mariotti, et al. (2005). "Terrestrial mechanisms of interannual CO ₂ variability,						
14								
15		Terrestrial CO ₂ sink as a residual	Models					
16	Year	of the global carbon budget	CLM4CN	HYLAND	LPJ-GUESS	LPJ		
17	1959	0,42	0,79	2,02	0,42	-0,83		
18	1960	1,14	0,75	1,53	1,16	0,81		
19	1961	1,20	0,30	1,71	-0,07	-0,55		
20	1962	1,76	0,79	2,37	1,25	0,57		
21	1963	1,72	-1,20	1,81	0,26	-0,37		

The Importance of Handover

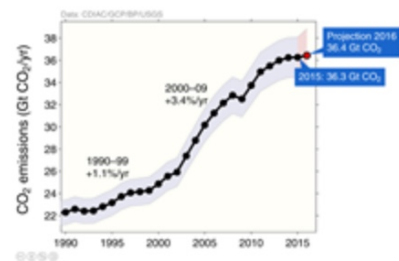
cdiac.ess-dive.lbl.gov/GCP/carbonbudget/2015/FWD via OSTI (DOE)
Ursprünglich am Oak Ridge National Laboratory.

NOTICE (September 2017): Data previously stored at the CDIAC archive are being transitioned to the new, ESS-DIVE archive. This website provides access to the CDAIC data during the transition. If you have any questions regarding the new archive or the data transition, please contact ess-dive-support@lbl.gov.

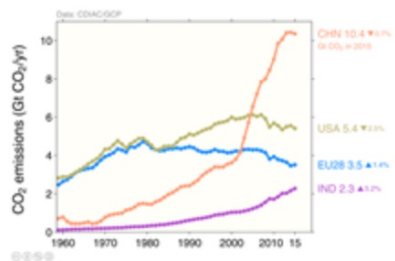


Global Carbon Budget 2016

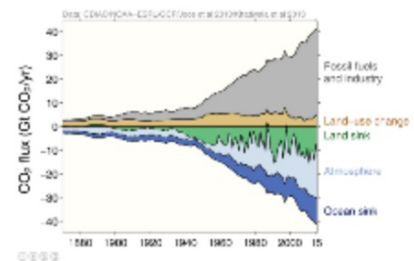
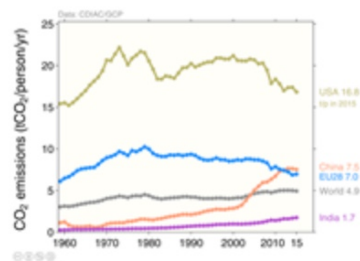
DOI: 10.5194/essd-8-605-2016



[2016 Global Budget v1.0](#)
(November 2016)

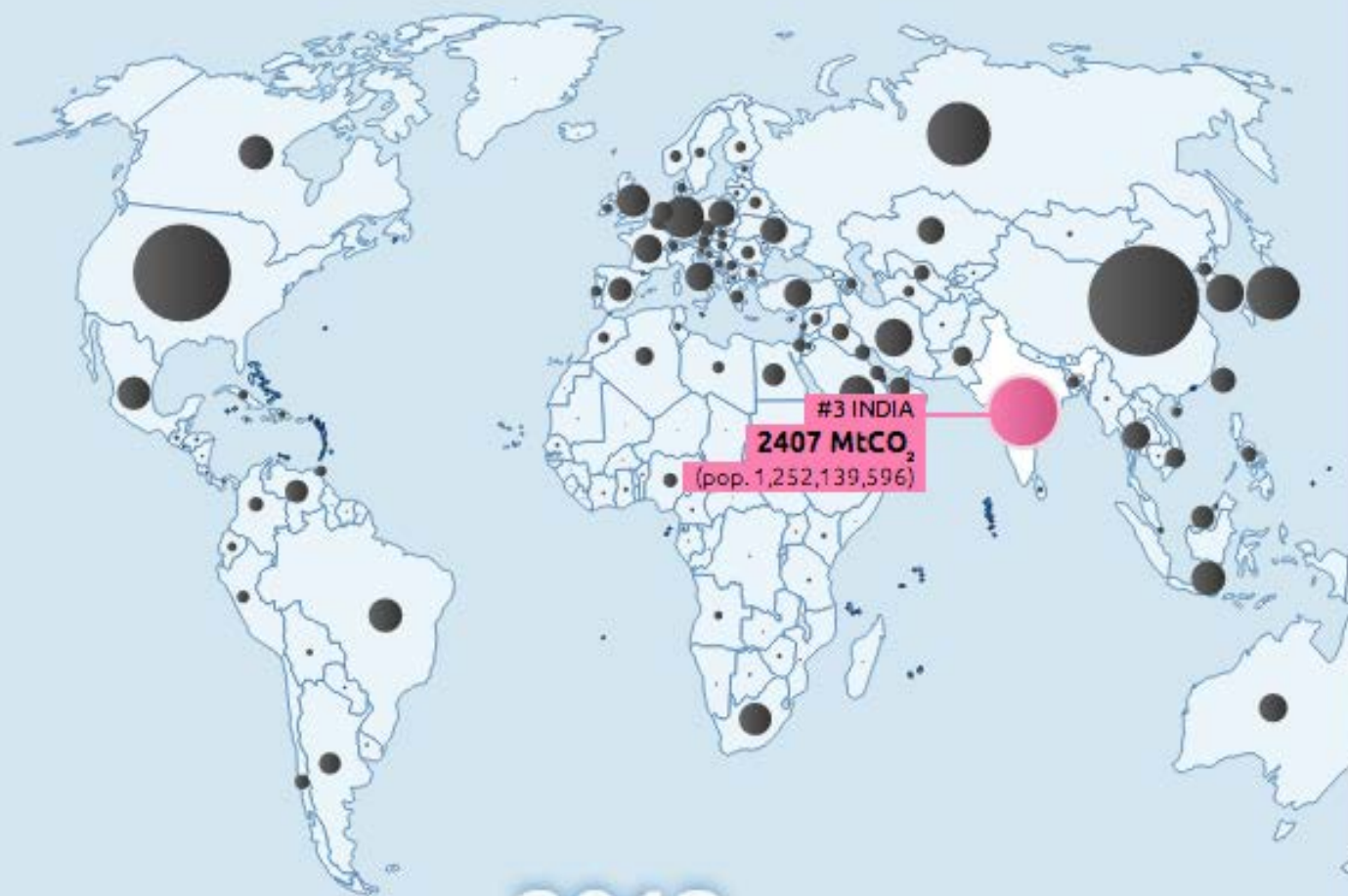


[National Emissions v1.0](#)
(November 2016)



Global Carbon Atlas - Visualization

- Note the **two panels** in www.globalcarbonatlas.org
 - **data journalism** (“emissions”)
 - **scientific visualization** (“research”)
- At the project website www.globalcarbonproject.org “real life” and **linkage between publications and data**



2013

World Total: 36131 MtCO₂

TOOLS

Fossil Fuels Emissions

EMISSIONS:
Territorial

UNIT:
MtCO₂

COUNTRIES:
Africa (54)
Asia (36)
Central America (31)
Europe (43)
Middle East (15)
North America (5)
Oceania (17)
South America (14)



MAP VIEW



CHART VIEW



FOCUS



RANKING



TIME SERIES



BUBBLES

SOURCES

HELP

METHODS

SHARE

DOWNLOAD



2013

Conclusions

- „Earth Science“ is a „Big Data“ problem (“Variety” in 3/5 Vs)
 - finding and exploiting patterns in metadata and data
- **Still needs Publishing processes**
 - For **quality** assurance and **recognition**
 - Published text is the best „**metadata**“ one can have
 - Articles are *still* the **linking hubs** in the **digital assets ecosystem**
- Needs **trustworthy infrastructures for data, software, ...**
 - most of all **skilled people** doing all the curatorial stuff etc.
- **Clever systems exploiting all this for discovery, aggregation, analysis, ... real time alerting, disaster mitigation, ...**

Thank you!

earth-syst-sci-data.net

expedition.awi.de

os.helmholtz.de

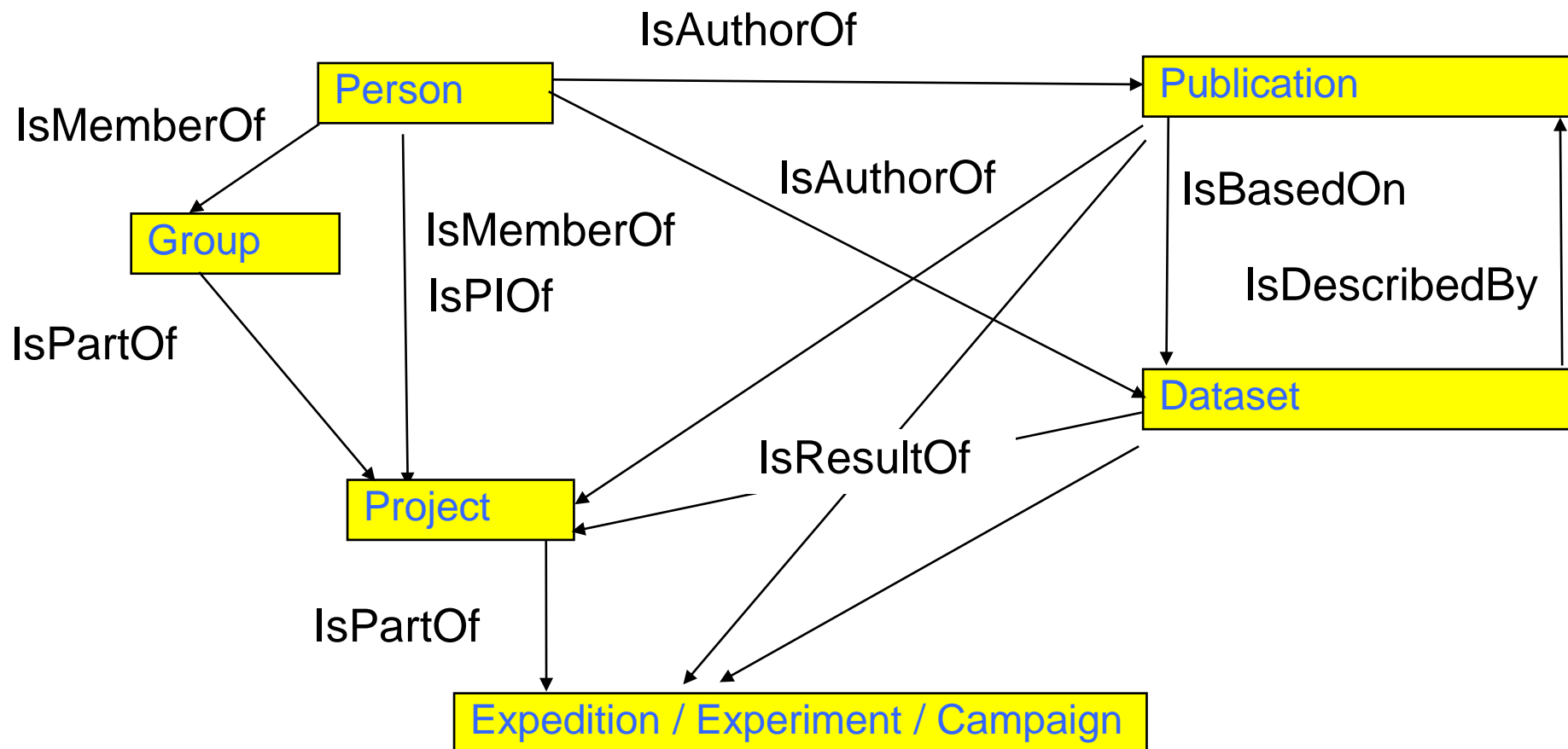
copdess.org

Bonus Material
- (almost) not shown in session

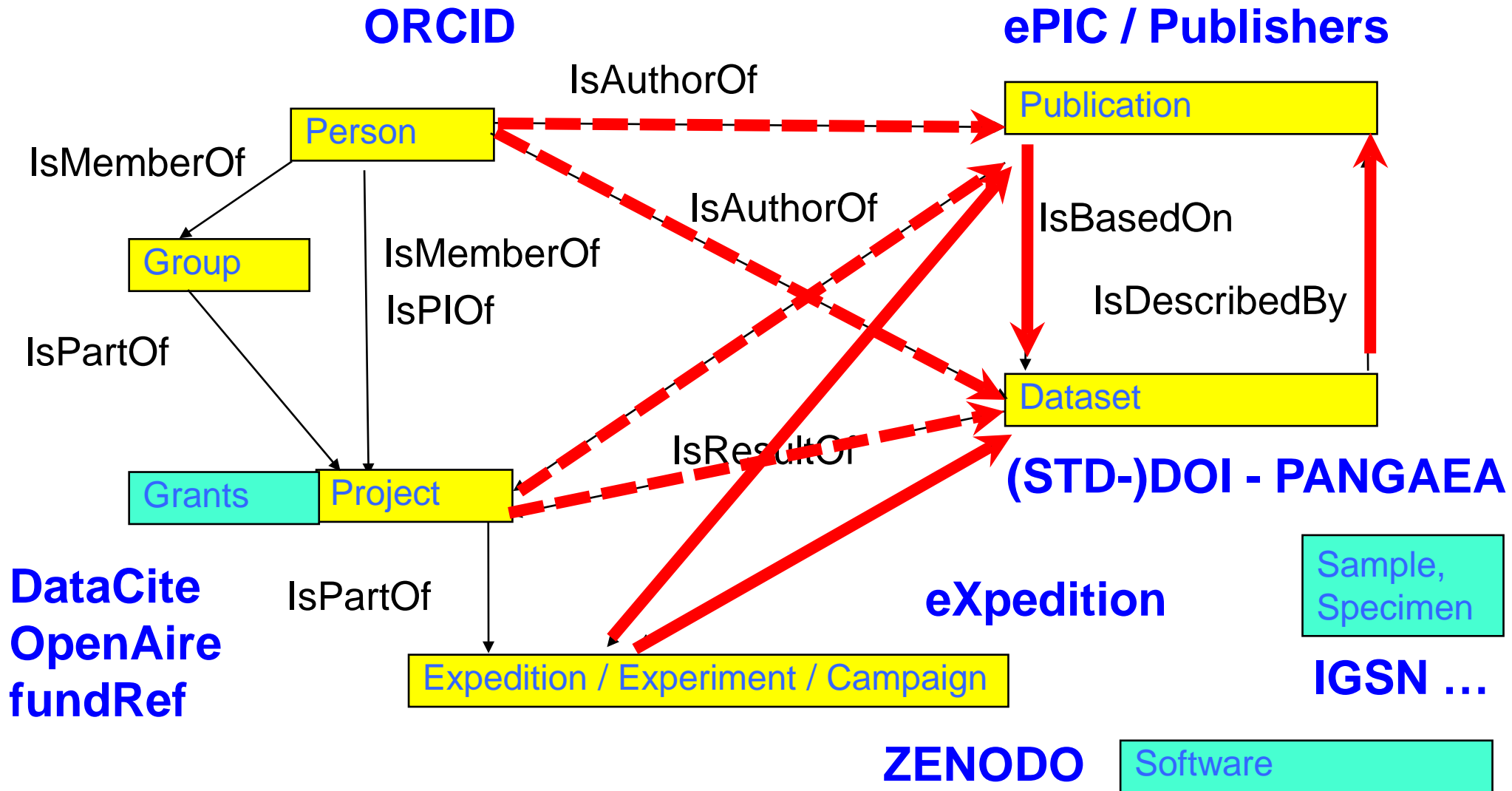
There is more than separate
publications and data!

Let's link it!

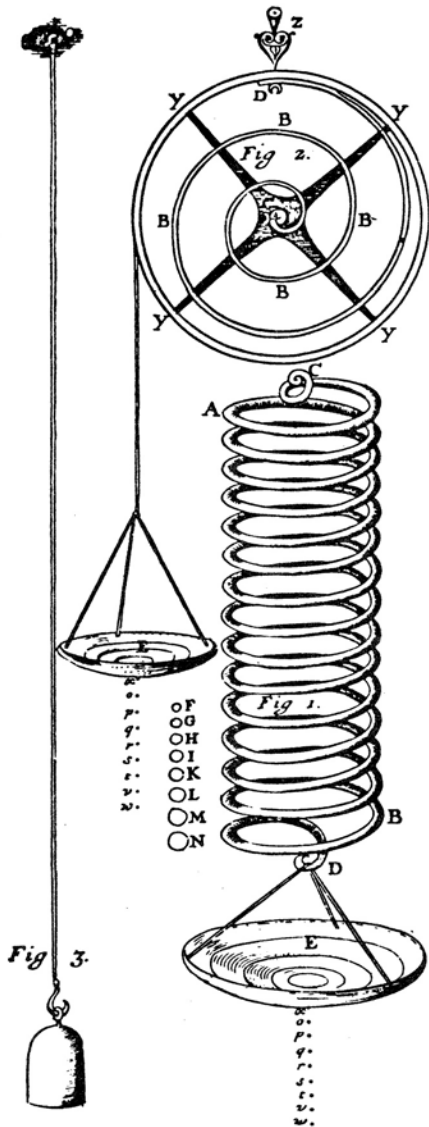
Pfeiffenberger, Macario, Text, Data and People, OAI4, CERN 2005



Pfeiffenberger, Carlson, TR32 DM WS Cologne 2014



Scared in the 17th Century



Hooke, published his law

1676 by anagram „ceiinosssstuv“

1678 in booklet

PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol I.
For Anno 1665, and 1666.

In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little with-
out Temple-Bar, and James Allestry in Duck-Lane,
Printers to the Royal Society.

PANGAEA – Elsevier

 [Purchase PDF \(743 K\)](#) |  [Export citation](#)

Abstract | [Article](#) | [Figures/Tables](#) | [References](#)


Marine Micropaleontology

Volume 66, Issues 3-4, 20 February 2008, Pages 192-207




doi:10.1016/j.marmicro.2007.09.002 | [How to Cite or Link Using DOI](#)

Copyright © 2007 Elsevier B.V. All rights reserved.

 [Permissions & Reprints](#)

 [Cited By in Scopus \(2\)](#)

Organic matter rain rates, oxygen availability, and vital effects from benthic foraminiferal $\delta^{13}\text{C}$ in the historic Skagerrak, North Sea

Sylvia Brückner  ^a,  and Andreas Mackensen ^a, 

^aAlfred Wegener Institute for Polar and Marine Research, Columbusstr., D-27568 Bremerhaven, Germany

Received 27 March 2007; revised 21 September 2007; accepted 24 September 2007. Available online 4 October 2007.

Abstract

The sediment cores 225514 and 225510 were recovered from 420 and 285 m water depth, respectively. They were investigated for their benthic foraminiferal $\delta^{13}\text{C}$ during the last 500 years.

[Purchase the full-text article](#) 

-  PDF and HTML
-  All references
-  All images
-  All tables



PANGAEA® – Supplementary Data

Stable carbon isotope composition of benthic foraminifera from sediments of the Skagerrak...



Related Articles

- [The tropical rainbelt and productivity changes off north...](#)
Marine Micropaleontology
- [Temporal variability in living deep-sea benthic foramin...](#)
Earth-Science Reviews
- [Early Maastrichtian benthic foraminiferal assemblages f...](#)
Marine Micropaleontology

eXpedition (in production since 2005)

Related Information: ["Reports on Polar and Marine Research"](#) (1982 to date)

[Primary data](#) (all polarstern datasets in PANGAEA)

[Handbook and scientific device documentation](#)(in deutsch)

[DShip](#) (Polarstern Data Acquisition System)

[VirtualPS: Virtual Polarstern Tour](#)

Expedition	Date Port	Region Research	Publications & Primary Data	Details
ANT-XXI/3 Coordinator: Pörtner, H. Chief scientist: Smetacek, V.	21.01.2004 - 25.03.2004 Capetown - Capetown [Map(png)]	Atlantic/Indian Ocean, Polar frontal zone Biology, EIFEX	ePIC: Publications ePIC: Reports on Polar and Marine Research ePIC: Weekly reports PANGAEA: Stations PANGAEA: Datasets <i>[Note: Publications and datasets for recent cruises may not yet be available]</i> Meteorology	▶
ANT-XXI/4	27.03.2004 - 06.05.2004	Lazarev Sea Biology, Krill, GLOBEC	ePIC: Publications ePIC: Reports on Polar and Marine Research	▶



Latest expeditions

PS88.2 (Polarstern)
Begin 2014-11-03 - Las Palmas de Gran Canaria
End 2014-11-29 - Cape Town

HE437 (Heincke)
Begin 2014-11-25 - Bremerhaven
End 2014-11-29 - Bremerhaven

Platforms

- Heincke
- Polarstern

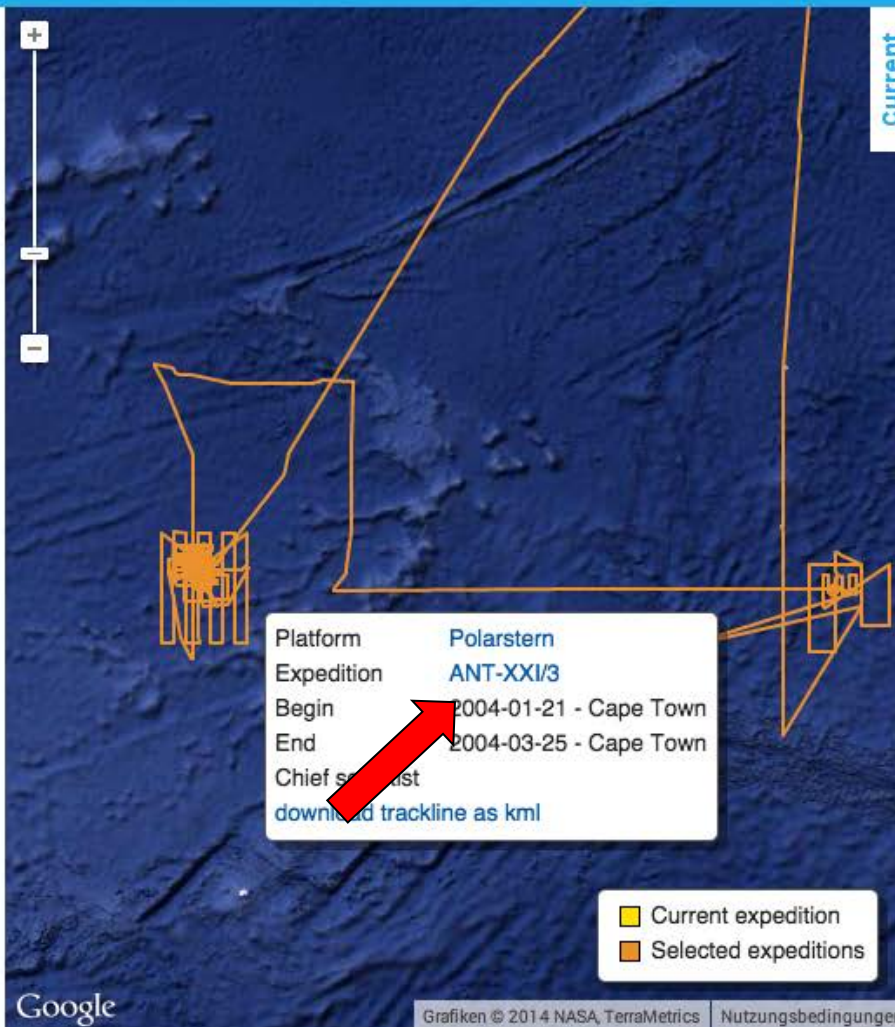
Temporal coverage

Begin
End

Search for keyword

Layers

- Stations
- DTM Framstrait
- DTM Knipovich Ridge
- DTM Scotia Sea
- Arctic Sea Ice
- Antarctic Sea Ice



Current

▶ Near real time data

▶ Latest data and publications

Archived data

- Continuous liquid water path (LWP) and integrated water vapour (IWV) measurements during POLARSTERN cruise PS83 (ANT-XXIX/10) on 2014-04-11 (2014)
- Continuous liquid water path (LWP) and integrated water vapour (IWV) measurements during POLARSTERN cruise PS83 (ANT-XXIX/10) on 2014-04-10 (2014)
- Continuous liquid water path (LWP) and integrated water vapour (IWV) measurements during POLARSTERN cruise PS83 (ANT-XXIX/10) on 2014-04-09 (2014)

Publications

- New insights into the past glaciation of the northeast Greenland continental shelf (2014)
- Taxon-specific epibenthic foraminiferal $\delta^{18}O$ in the Arctic Ocean: relationship to water masses, deep circulation, and brine release (2014)
- Geo-Morphological Analyses of the Gakkel Ridge and the Southwest Indian Ridge (2014)

Reports

- Expedition Programme PS88 (2014)
- Expeditionsprogramm Polarstern Nr. 93 PS84, PS85, PS86, PS87 (ARK-XXVIII/1-4), (2014)
- Effects of cold glacier ice crystal anisotropy on seismic data (2014)

Bookmark your search

Provider All Archived data **Publications** Reports

EPIC (114)

Near real time

AWI

Sort by

best match date

Keywords



Region
(no facets available)

Author

- Peeken, Ilka (34)
- Bracher, Astrid (33)
- Smetacek, Victor (25)
- Assmy, Philipp (23)

Intercomparison of ocean color products identifying coccolithophore blooms on global and regional scales. (2012) +
Bracher, Astrid; Sadeghi, Alireza; Dinter, Tilman; Altenburg Soppa, Mariana; Peeken, Ilka; Rozanov, Vladimir; Taylor, Bettina; Vountas, Marco
hdl:10013/epic.40317

+iron +response

Search

Bookmark your search

Provider

EPIC (11)

All Archived data **Publications**

AWI

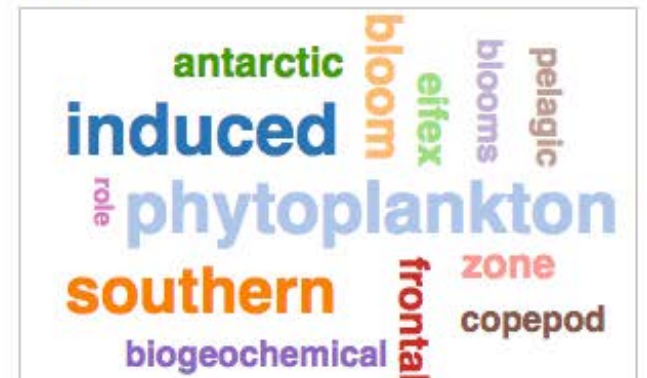
Sort by

best match

Reports Near real time

date

Keywords



Region
(no facets available)

Author

- Bathmann, Ulrich (6)
- Kräßelsky, Sören (5)
- Assmy, Philipp (4)
- Henies, Joachim (4)

Selected facets: 3 x

Ecological and biogeochemical response of Antarctic ecosystems on iron fertilization and implications on global carbon cycle (2004) +
Bathmann, Ulrich
hdl:10013/epic.22812



Bookmark your search

Provider: PANGAEA (267)

Archived data Publications Reports

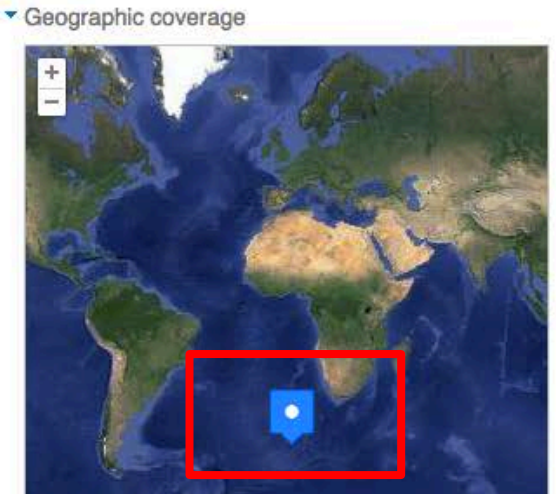
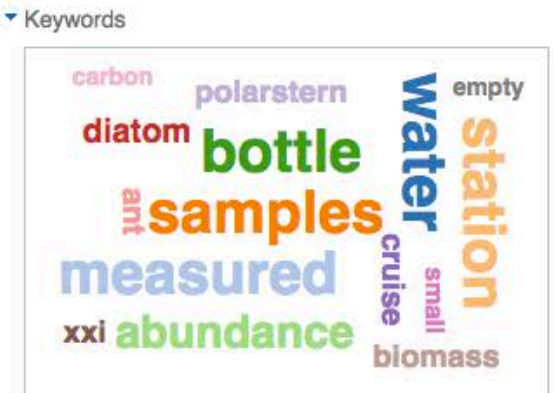
Near real time

AWI

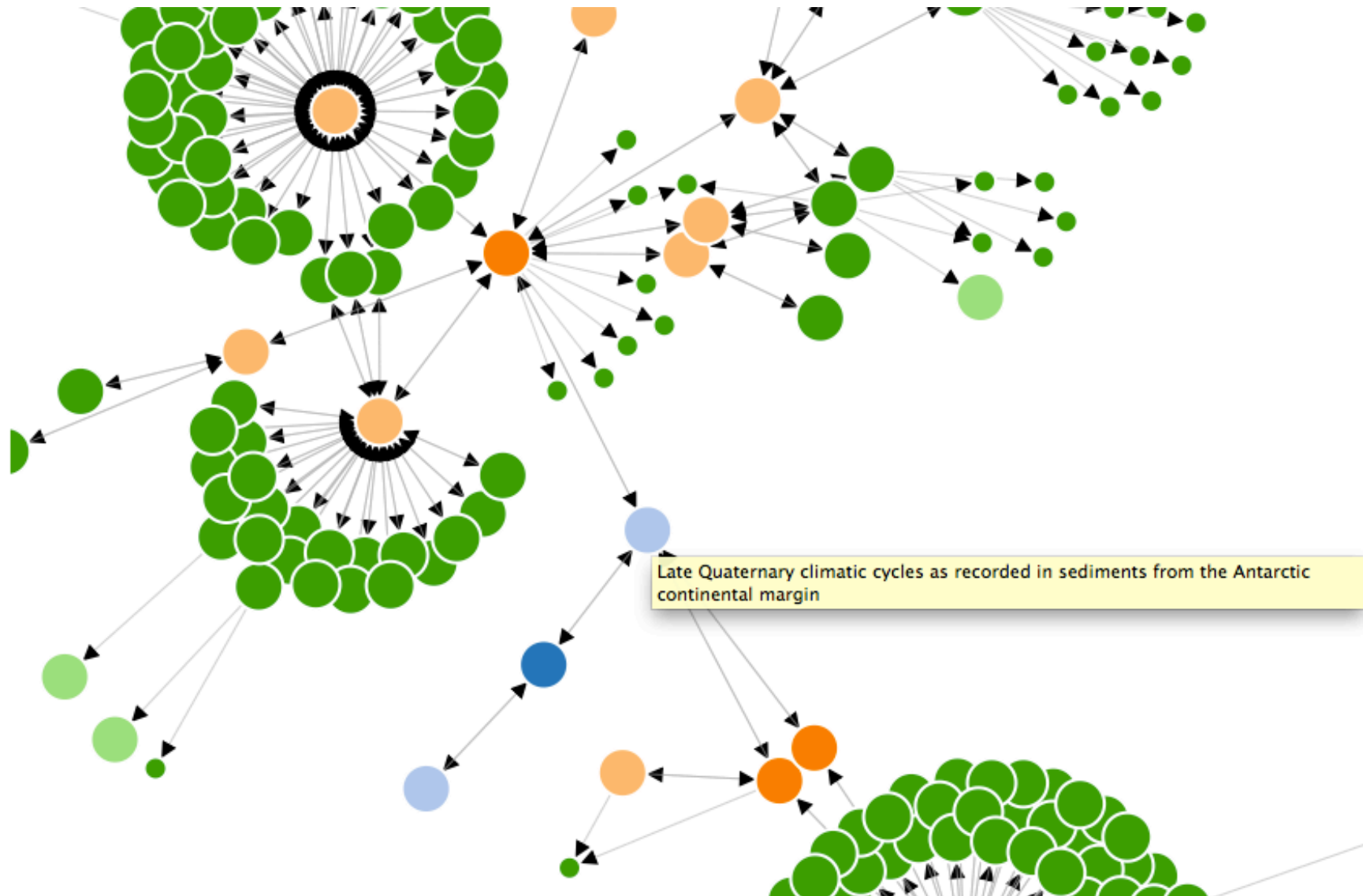
Sort by: best match date

- Region
 - Atlantic Ocean (267)
 - Indian Ocean (1)
- Author
 - Assmy, Philipp (129)
 - Henjes, Joachim (68)
 - König-Langlo, Gert (65)
 - Cortese, Giuseppe (2)
 - Gersonde, Rainer (2)
 - more facets...
- Platform
 - Polarstern (267)
 - Joides Resolution (1)
- Expedition
 - ANT-XXI/3 (267)
 - ANT-VI/3 (1)
 - ANT-XI/2 (1)
 - ANT-XVIII/2 (1)
 - Leg177 (1)
- Device
 - CTD/Rosette (100)

- Selected facets: 3
- Microzooplankton biomass measured on water bottle samples at station PS65/427-1 (2007)
Henjes, Joachim
doi:10.1594/PANGAEA.603272
[tsv](#) [view object](#)
 - Microzooplankton biomass measured on water bottle samples at station PS65/470-1 (2007)
Henjes, Joachim
doi:10.1594/PANGAEA.603274
[tsv](#) [view object](#)
 - Microzooplankton biomass measured on water bottle samples at station PS65/514-2 (2007)
Henjes, Joachim
doi:10.1594/PANGAEA.603276
[tsv](#) [view object](#)
 - Microzooplankton biomass measured on water bottle samples at station PS65/545-1 (2007)
Henjes, Joachim
doi:10.1594/PANGAEA.603278
[tsv](#) [view object](#)



eXpedition – Publications and Data network



Impressions from COPDESS

**- Coalition on Publishing Data in the
Earth and Space Sciences -**

JOINT DECLARATION OF DATA CITATION PRINCIPLES - FINAL

When citing please use: Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014
[<https://www.force11.org/datacitation>].

ENDORSEMENT LIST

DC¹

Data Citation Principles

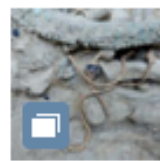
PREAMBLE

Sound, reproducible scholarship rests upon a foundation of robust, accessible data. For this to be so in practice as well as theory, data must be accorded due importance in the practice of scholarship and in the enduring scholarly record. In other words, data should be considered legitimate, citable products of research.

Data citation, like the citation of other evidence and sources, is good research practice and

Journals' Transparency Criteria

- TOP: modular, agnostic to disciplines, low barrier to entry; categories are:
 - a) **data citation**
 - b) design transparency (**standard operating procedures**, protocols)
 - c) materials transparency
 - d) **data transparency**
 - e) analytical methods (code) transparency
 - f) preregistration of studies
 - **g) preregistration of analysis plans**
 - h) replication
- 3 levels
 - level 1: article states whether data are available and where to access
 - level 2: data has to be in trusted repository
 - level 3: data in TR, but reported analyses will be reproduced independently



Grave of 'Griffin Warrior' at Pylos Could Be a Gateway to Civilizations



Deadly Heat Is Forecast in Persian Gulf by 2100



MATTER DNA of Ancient Children Offers Clues on How People Settled the...

Journal
• > 526 j

SCIENCE

Journal Science Releases Guidelines for Publishing Scientific Studies

By BENEDICT CAREY JUNE 25, 2015

- Email
- Share
- Tweet
- Save
- More

In the midst of a debate over scientific misconduct, one of the world's leading scientific journals on Thursday posted the most comprehensive guidelines for the publication of studies in basic science to date, calling for the adoption of clearly defined rules on the sharing of data and methods.



Dr. Marcia McNutt, the editor-in-chief of Science, co-authored new guidelines to improve transparency of data, methods and materials in published studies. Drew Angerer for The New York Times



The guidelines, published in Science, come weeks after the

Directory of trustworthy repositories (for ESS):

- **Discipline-specific, because only those**
 - can catch the needed metadata
 - can be regarded sustainable
- **In order of specificity:**
 - re3data.org
 - „something" developed by AGU/NSF/COS
 - journals' lists (ESSD, Scientific Data)
- **Certification**
 - WDS (actually an accreditation process)
 - Data Seal of Approval

Identifiers

- **For data: persistent identifiers for data,**
 - ⇒ preferably DOIs, implying **fixity, integrity!**
- **For authors, contributors: ORCID**
 - ⇒ all publishers to use ORCID, even the ORCID AuthN service
- **For samples: IGSN**
- **For grants: FundRef (<http://www.crossref.org/fundref/>)**
 - ⇒ <http://dx.doi.org/10.13039/501100000780>,
"European Commission",
 - "narrower": [{"resource":
"http://dx.doi.org/10.13039/501100000889"}],

Transparency = Openness?

- **The “How” of Open**

- Is registration acceptable ??
- Which kind of license (if any)

- **The “When” of Open**

- Some time after end of data acquisition?
- Some time after end of project / funding?
- **At time of publication?!**
 - After acceptance?
 - If there is an on-going embargo, how can reviewers have access?