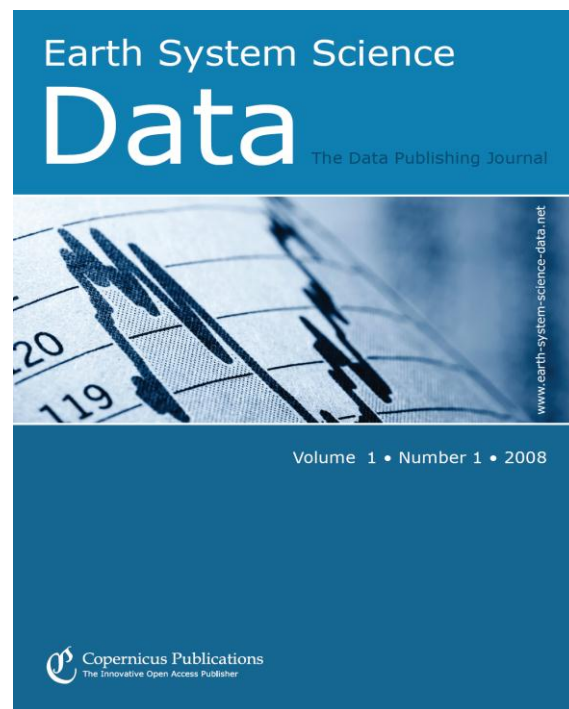


Repositorien für Text und Forschungsdaten

Hans Pfeiffenberger

Alfred-Wegener-Institute for Polar and Marine Research,
Helmholtz Association - Germany

Repositorien Praxis und Vision 2010-11-30, Berlin



AWI 

Agenda

- **Das Umfeld**
 - Riding the Wave
 - Allianz / KII
 - Verlage und Zeitschriften
- **Beispiel(e) aus der realen Welt**
 - Ein soziologisches Problem?
 - Eine Vision (von 2005)
- **Was tun? Dienst an der Wissenschaft !**
 - ... über deren Communities
 - ... Akzeptanz, Nutzung erzielen

Making this a reality is a more difficult task than it may seem. To collect, curate, preserve and make available ever-increasing amounts of scientific data, new types of infrastructures will be needed. The potential **benefits are enormous but the same is true for the costs.**

Neelie Kroes, VP European Commission

Riding the wave

How Europe can gain from the rising tide of scientific data

Final report of the High Level Expert Group on Scientific Data
A submission to the European Commission

October 2010



Salvatore Mele, APA2010:

“Preservation, re-use and access” are exciting, are they?

Wissenschaftsorganisationen und Förderer in Deutschland

- **Allianz der Wissenschaftsorganisationen**
 - Grundsätze zum Umgang mit Forschungsdaten (Juni 2010)
 - „Entwicklung von Infrastrukturen: **Ein nachhaltiges Forschungsdaten-Management stellt vielfältige technische und organisatorische Anforderungen.** ... Infrastrukturen sind gemäß diesen Anforderungen zu entwickeln und möglichst von Beginn **an in internationale und interdisziplinäre Netzwerke interoperabel einzubinden.**
- **Wissenschaftsministerkonferenz**
 - Kommission zukünftige Informationsinfrastruktur (2011)
 - Konkretisierungen ...
 - **Strukturen, Aufwandsschätzung ???**

www.nature.com/nature

Data's shameful neg

Research cannot flourish if data are not preserved and

Zitierungen steigen, wenn
Data-Supplement vorhanden...

Aber wer hostet es?

Soll das Supplement bei Peer
Review mit begutachtet
werden?

Neue Geschäftschance!!

and access to digital data are central to their mission, and need to be supported accordingly. Organizations in the United Kingdom for instance, have made a good start. The Joint Information Systems

NATURE INSIGHT TRANSCRIBING THE GENOME

10 September 2009 | www.nature.com/nature | £10 THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

nature

NATUREJOBS
Philadelphia
calling



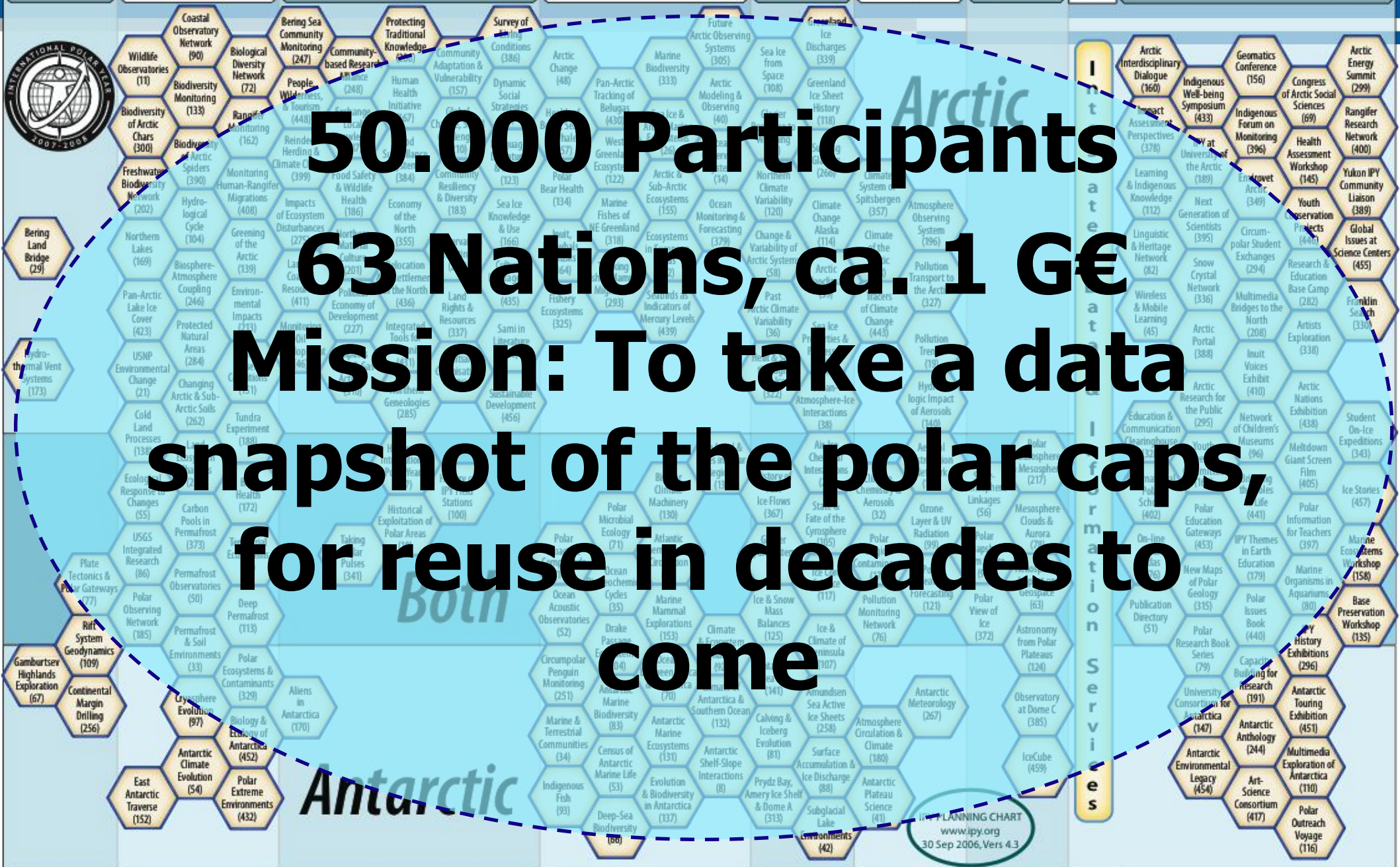
DATA — WHAT DATA?
Learning to share your results

LUNAR EXPLORATION
Highland games

THE HUMAN BRAIN
Procrastination pathways

VERTEBRATE EVOLUTION
What jawless fish say about us

9 770028 083095 37 >



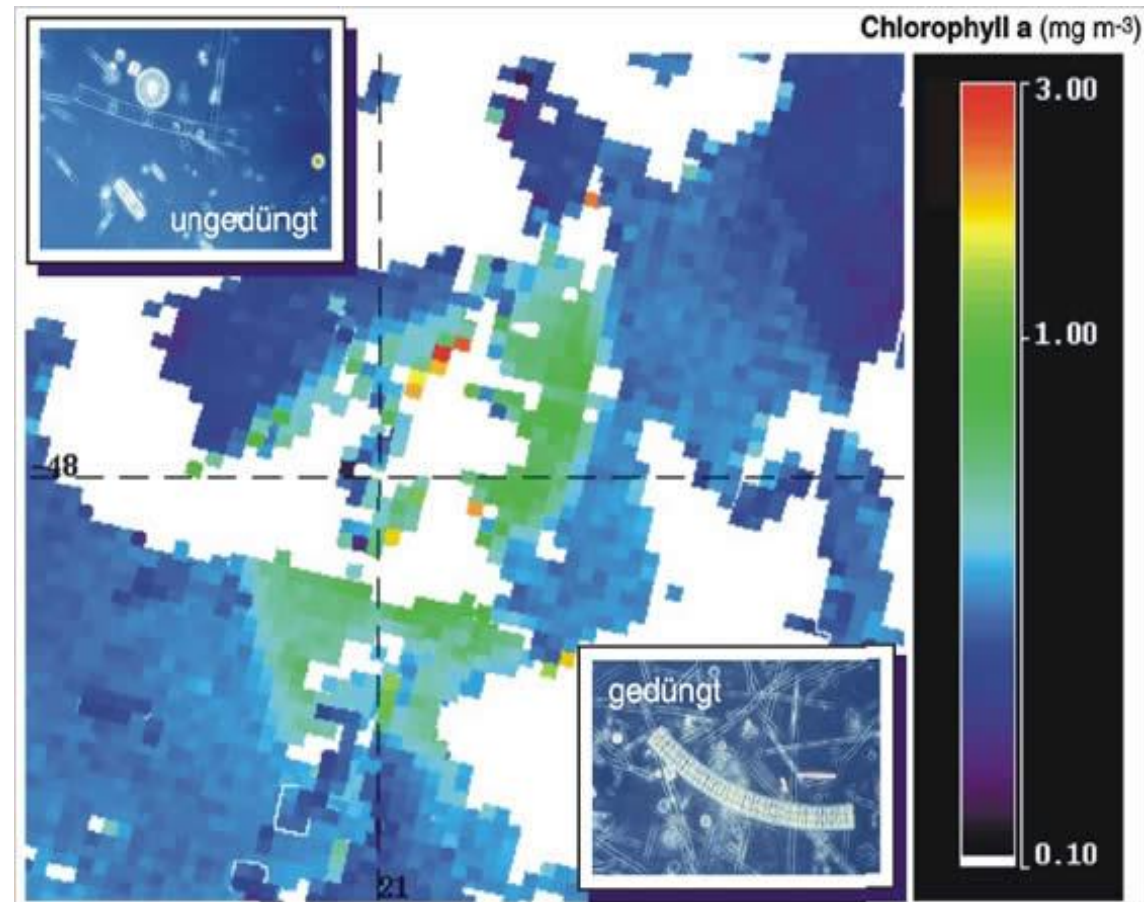
50.000 Participants
63 Nations, ca. 1 G€
Mission: To take a data snapshot of the polar caps, for reuse in decades to come




IPY PLANNING CHART
www.ipy.org
30 Sep 2006, Vers 4.3

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 European countries and South Africa
 - **Oceanographers**
 - **Biologists**
 - **Chemists.....**
- **“Biogeochemistry”**



eXpedition (Alpha)



Select research platform, year, campaign
 Polarstern | 2004 | ANT-XXI/3

Add to trackline map:

- Trackline for ANT-XXI/3
- PANGAEA: along track
- PANGAEA: stations
- PANGAEA: Radiosonde
- Publications

Information on ANT-XXI/3

Platform: R/V Polarstern
 Operation: AWI Bremerhaven
 Scientific devices: WIKI documentation
 Coordinator: Pörtner, H.
 Chief scientist: Smetacek, V.
 Region: Atlantic/Indian Ocean, Polar frontal zone
 Research: Biology
 Program: EIFEX
 Start: 2004-01-21, Cape Town
 End: 2004-03-25, Cape Town

Documents on ANT-XXI/3

[Cruise Summary Report](#)

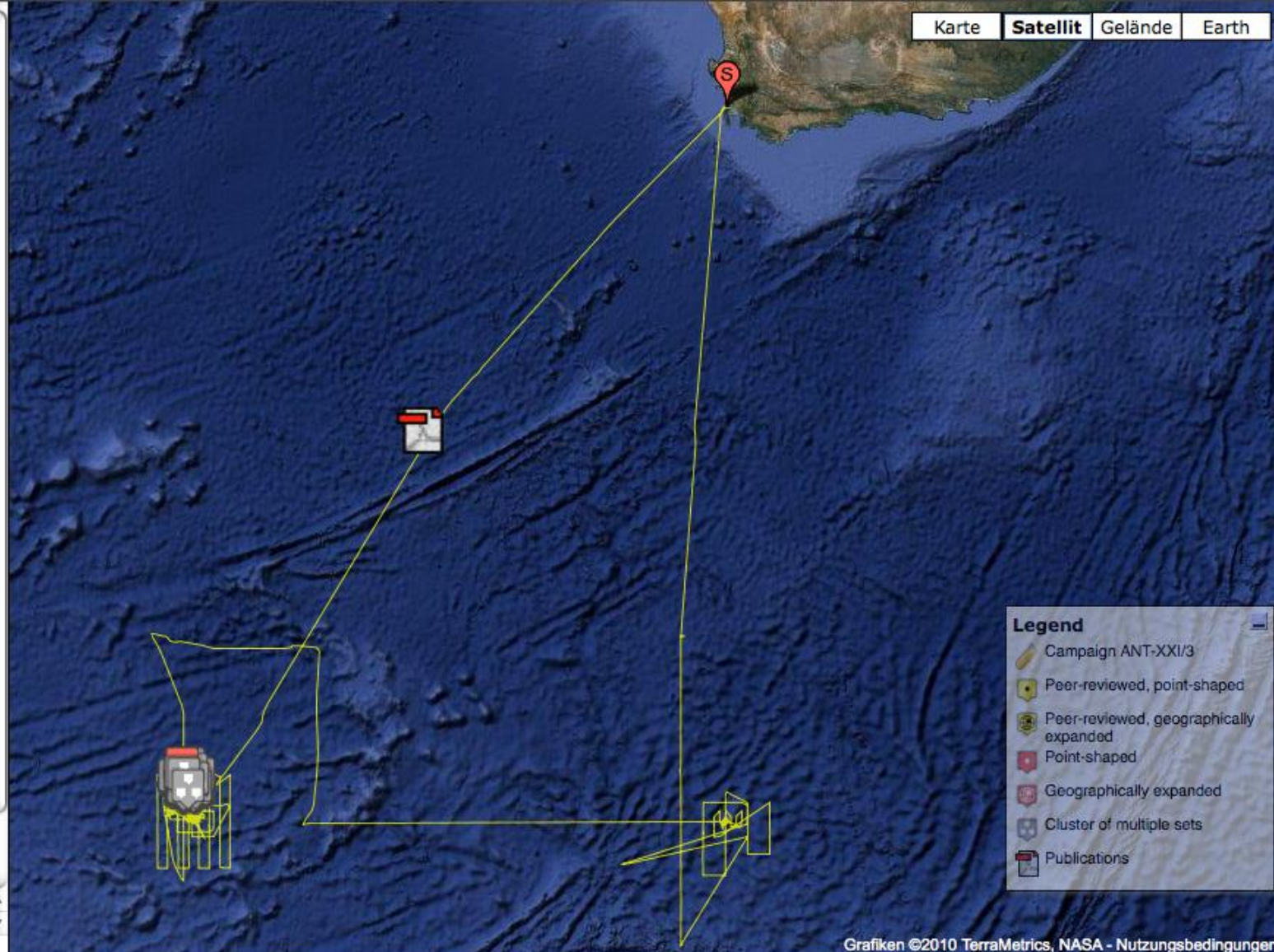
ePIC: [all Publications](#) | ["Berichte für Polar- und Meeresforschung"](#) | [Weekly reports](#)

Data on ANT-XXI/3

[DOMIN - Polarstern Data Acquisition System](#)
[PANGAEA: Datasets](#) | [Map](#) | [Stations](#) | [WIKI](#)

General Information

[Home](#) | [About eXpedition](#) | [Contact](#) | [Impressum](#)
[AWI stations](#)



PANGAEA EIFEX

94. **Henjes, J (2007):** Microzooplankton abundance measured on water bottle samples at station PS65/593-3

Size: 215 data points

doi:10.1594/PANGAEA.603343 - Score: 47% - Similar datasets

95. **Jansen, S (2005):** Egg production rate of *Rhincalanus gigas* from bongo net stations in the Southern Ocean

Reference: **Jansen, S; Klaas, C; Krägefsky, S et al. (2006):** Reproductive response of the copepod *Rhincalanus gigas* to an iron-induced phytoplankton bloom in the Southern Ocean. *Polar Biology*

Size: 54 data points

doi:10.1594/PANGAEA.321035 - Score: 47% - Similar datasets




96. **Assmy, P (2007):** Phytoplankton abundance measured on water bottle samples at station PS65/424-3

Size: 453 data points

doi:10.1594/PANGAEA.603388 - Score: 35% - Similar datasets

ePIC (AWI IR) S.Jansen 2006 (peer reviewed)

2006

- [Jansen, S., Klaas, C., Krägefsky, S., Harbou, L. von , Bathmann, U.\(2006\).Reproductive response of the copepod *Rhincalanus gigas* to an iron-induced phytoplankton bloom in the Southern Ocean, Polar biology, 29\(12\), 1039-1044, doi:10.1007/s00300-006-0147-0 .](#) 
[\[>\] Full-text](#)
- [Jansen, S., Wexels Riser, C., Wassmann, P., Bathmann, U.\(2006\).Copepod feeding behaviour and egg production during a dinoflagellate bloom in the North Sea, Harmful Algae, 5, 102-112, doi:10.1016/j.hal.2005.06.006 .](#) 
[\[>\] Full-text](#)
- [Jansen, S.\(2006\).Feeding behaviour of calanoid copepods and analyses of their faecal pellets, PhD Thesis, 154 S. {http://elib.suub.uni-bremen.de/diss/docs/00010366.pdf }, Universität Bremen.](#) 
[\[>\] Full-text](#)

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](https://doi.org/10.1016/j.marmicro.2007.09.002) | [How to Cite or Link Using DOI](#)
 Copyright © 2007 Elsevier B.V. All rights reserved.

 Cited By in Scopus (2)

[Permissions & Reprints](#)

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

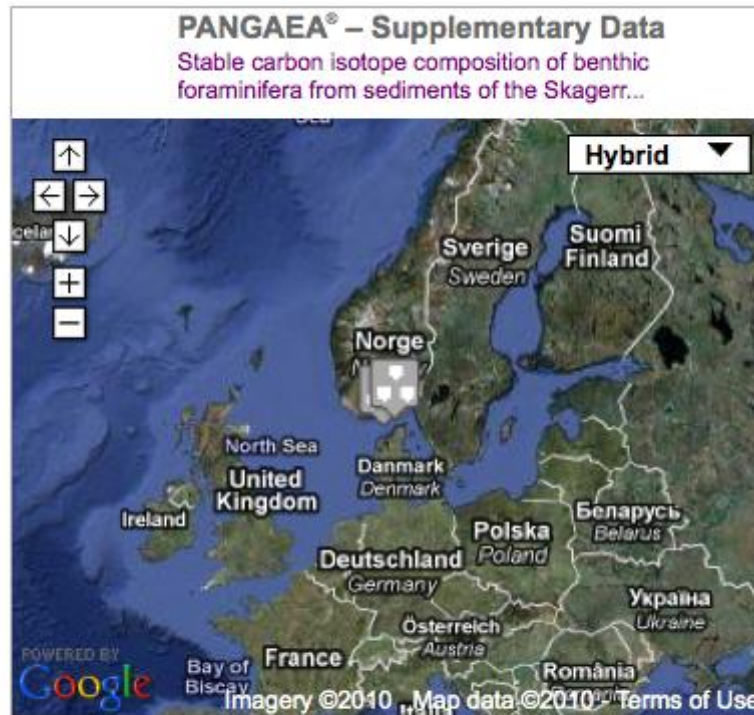
Received 27 March 2007; revised 21 September 2007; accepted 24 September 2007. Available online 4 October 2007.

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 

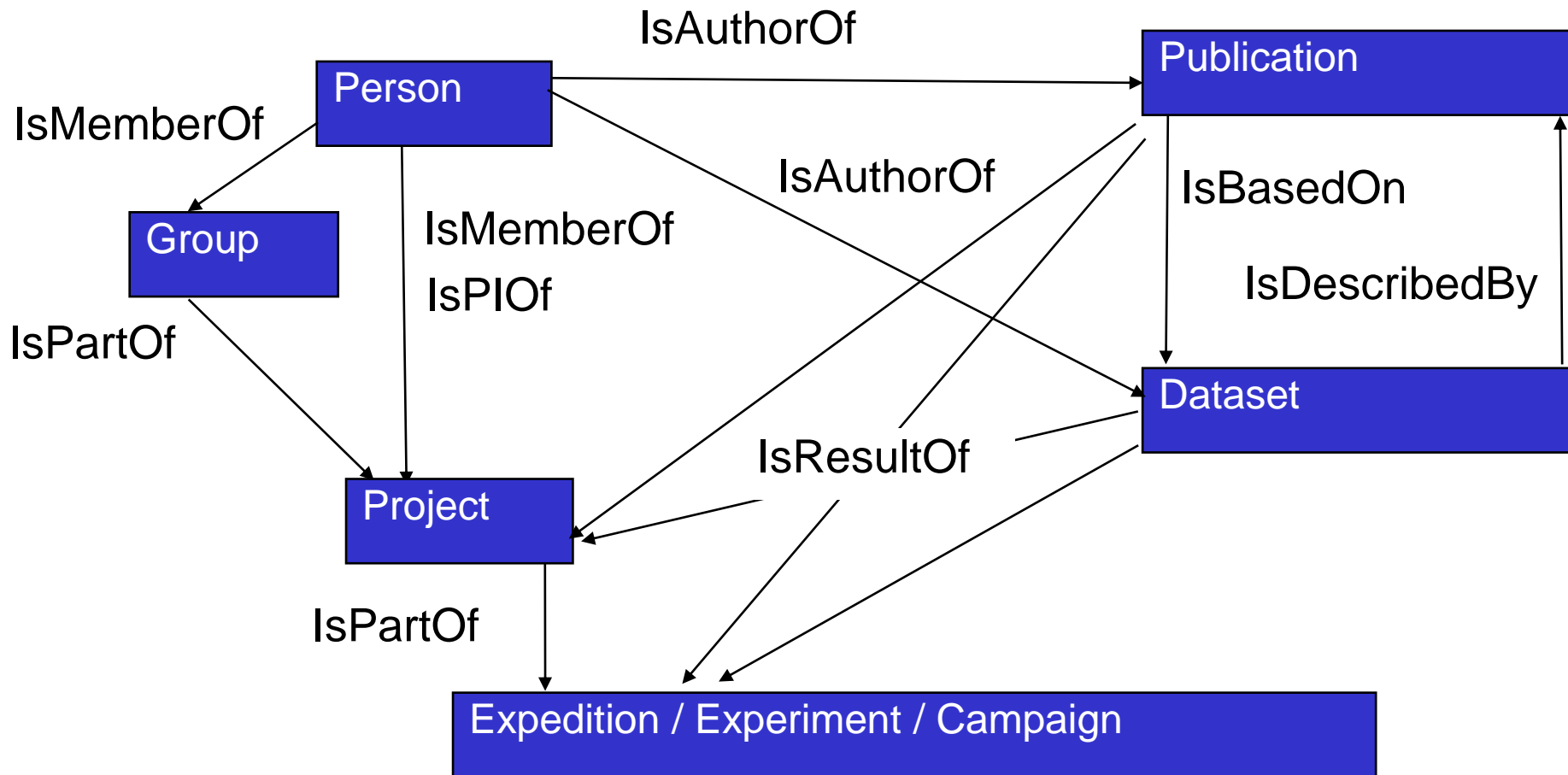
-  PDF and HTML
-  All references
-  All images
-  All tables



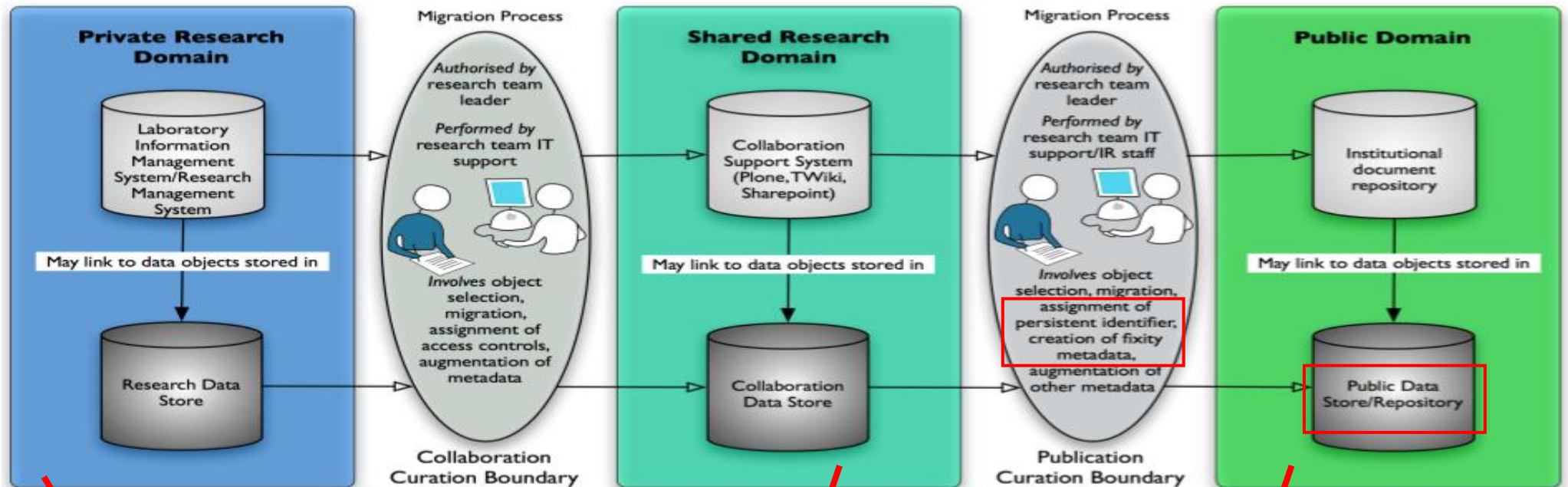
Related Articles

-  [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](#)

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



Landschaft des Informationsmanagements, Treloar 2007



This domain involves the core research team as they undertake the research, usually within a single institution. Access is often tightly controlled as hypotheses and analyses are developed.

This domain involves researchers outside the core team as they collaborate with colleagues, often across institutions. Access is more open, but not everything is shared.

This domain involves the public sphere (publication in the sense of making public). Access will usually be open to all.

Version 1.4, <http://andrew.treloar.net/>, 07Dec07

Domäne der 1980er DFG Empfehlungen zur Nachvollziehbarkeit ??

????

Institutional Repositories (?) Domäne der Langzeitarchivierung, Nachnutzung ??

How to identify a Scientific Community, match “user community”

- Organizes itself via a (learned) society
- Publishes in a range of journals **Share data via a range of repositories**
- Meets at a set of regular conferences
- Organizes large international research programmes

- Shares a vocabulary **Ontologies, thesauri, gazeteers**
- Agrees on a set of methods and/or practises **Codes to aggregate, validate, visualize**

To a limited extent

- Employs a certain instrument or research infrastructure
- ~~Contribute to a certain e-Infrastructure, software or hardware (“VO”)~~

Criteria for service adoption by communities

1 year horizon (Individual, piece of software)

- Usefulness (cost of learning/adoption vs. short term success)
- Availability (Access)
- Usability (No certificates! Classical GIS vs. Google Maps!)

10+ years horizon (Community, Infrastructure)

- **Trust** (Proven **scientific validity** / quality, conferred by branding)
- **Reliability** (Persistence, to become part of **scientific practise**)

Catch 22 (for service providers, visionaries)

- **Real big improvements for complex problems** (Earth System)
- **Trust builds slowly, bottom up** (per narrow discipline)