

Archiving data from earth system research

-

a show case



Hannes Grobe & Rainer Sieger

What is PANGAEA ?

Pangaea is an information system for **georeferenced data** from basic research on the earth system. Data are stored in a relational database which is accessible on the Internet. The system is open to individual scientists as well as projects to preserve data and to make it public available.

Intention and operation of Pangaea is comparable to a library - a public electronic library for research data.

The system faces any technical challenges, which a modern information system on the Internet should have. It is continuously developed and adopted to new technical and scientific requirements and standards.

WWW.PANGAEA.DE

Digital „Library“

in

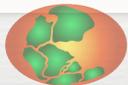
Open Access

for

Georeferenced Data

from

Earth & Environment



Does science need a *library* for data ?

MOTIVATION

- ✓ Open Access to scientific results
- ✓ Safeguarding 'Good Scientific Practice'
- ✓ Added value through integration
- ✓ Overview on 'what exists'
- ✓ Persistent identification and availability
- ✓ Reporting to funding organization
- ✓ Personal record + credit
- ✓ BackUp

Libraries preserve
the knowledge
of mankind!



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.

Initiatives & Protocols

Open Access

Budapest Open Access Initiative

 Deutsch  Français  На русском

[Home](#)

[Read the initiative](#)

[Sign on to the initiative](#)

[View signatures](#)

Budapest Open Access Initiative

The Budapest Open Access Initiative arises from a small but lively meeting convened in Budapest by the Open Society Institute (OSI) on December 1-2, 2001. The purpose of the meeting was to accelerate progress in the international effort to make research articles in all academic fields freely available on the

September 29, 2004

Grants for Open Access Journals
[Learn More](#)



MAX-PLANCK-GESELLSCHAFT

Program

OA Conference
Program Committee
ECHO Meeting

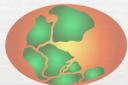
Conference on

Open Access to Knowledge in the Sciences and Humanities

20 - 22 Oct 2003, Berlin

Berlin Declaration

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities



**OECD Principles and
Guidelines for Access
to Research Data from
Public Funding**



Who are the hosts of PANGAEA ?

Pangaea is operated by the

(1) Alfred Wegener Institute for Polar and Marine Research,
member of the Helmholtz Association of National Research Centres,
funded by the Federal Ministry of Education and Research

and the

(2) Center for Marine Environmental Sciences at the University of Bremen
with support of the Department of Geoscience and the
research center ocean margins,
funded by the German Research Foundation.

Both institutions have committed to long-term operate **Pangaea** and the
World Data Center for Marine Environmental Sciences.

OPERATING INSTITUTIONS



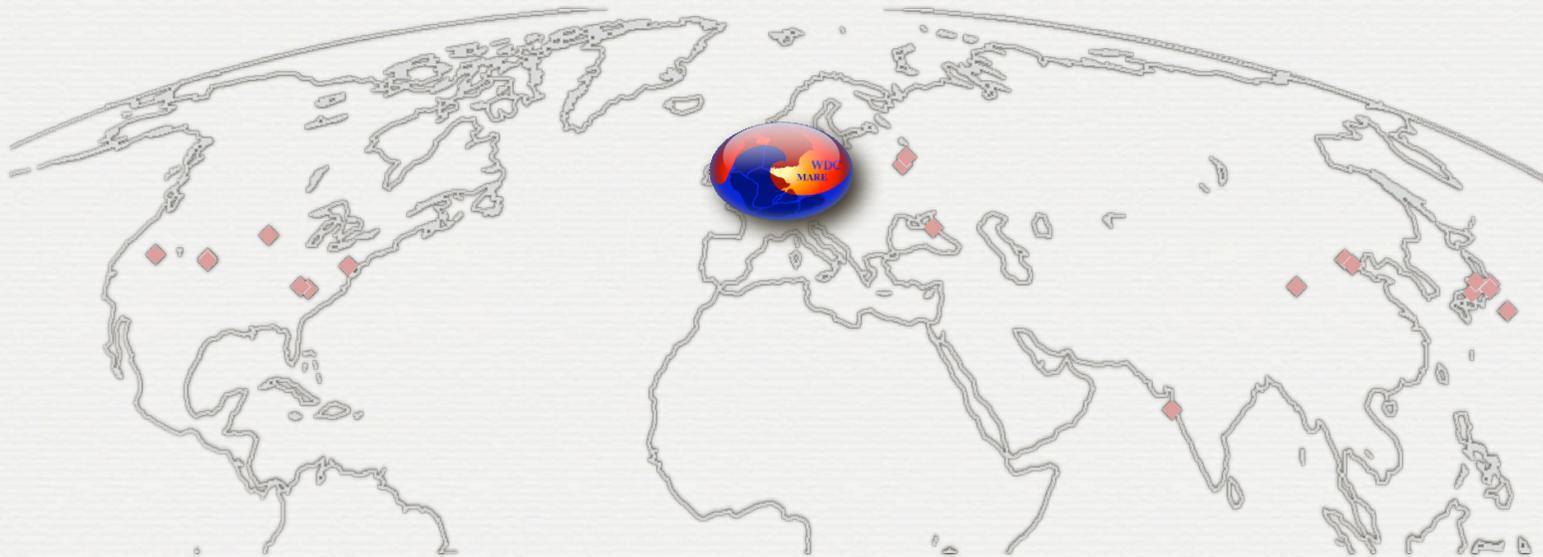
Center for Marine Environmental Sciences, Bremen



Alfred Wegener Institute for Polar and Marine Research
Bremerhaven



WORLD DATA CENTER SYSTEM OF ICSU



Final data report
for projects

CD/DVD with data and
local search engine

Description and further
information in a booklet

Distribution through 300
libraries with focus on
marine research

WDC-MARE 0001
Reports **2004**



Integrated Data Sets of the DFG Research Project SFB 313

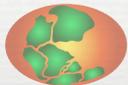
Environmental Change: The Northern North Atlantic
(Veränderungen der Umwelt: Der nördliche Nordatlantik)

Hannes Grobe, Michael Diepenbrock,
Priska Schäfer, Jörn Thiede & Gerold Wefer

WORLD DATA CENTER FOR MARINE ENVIRONMENTAL SCIENCES

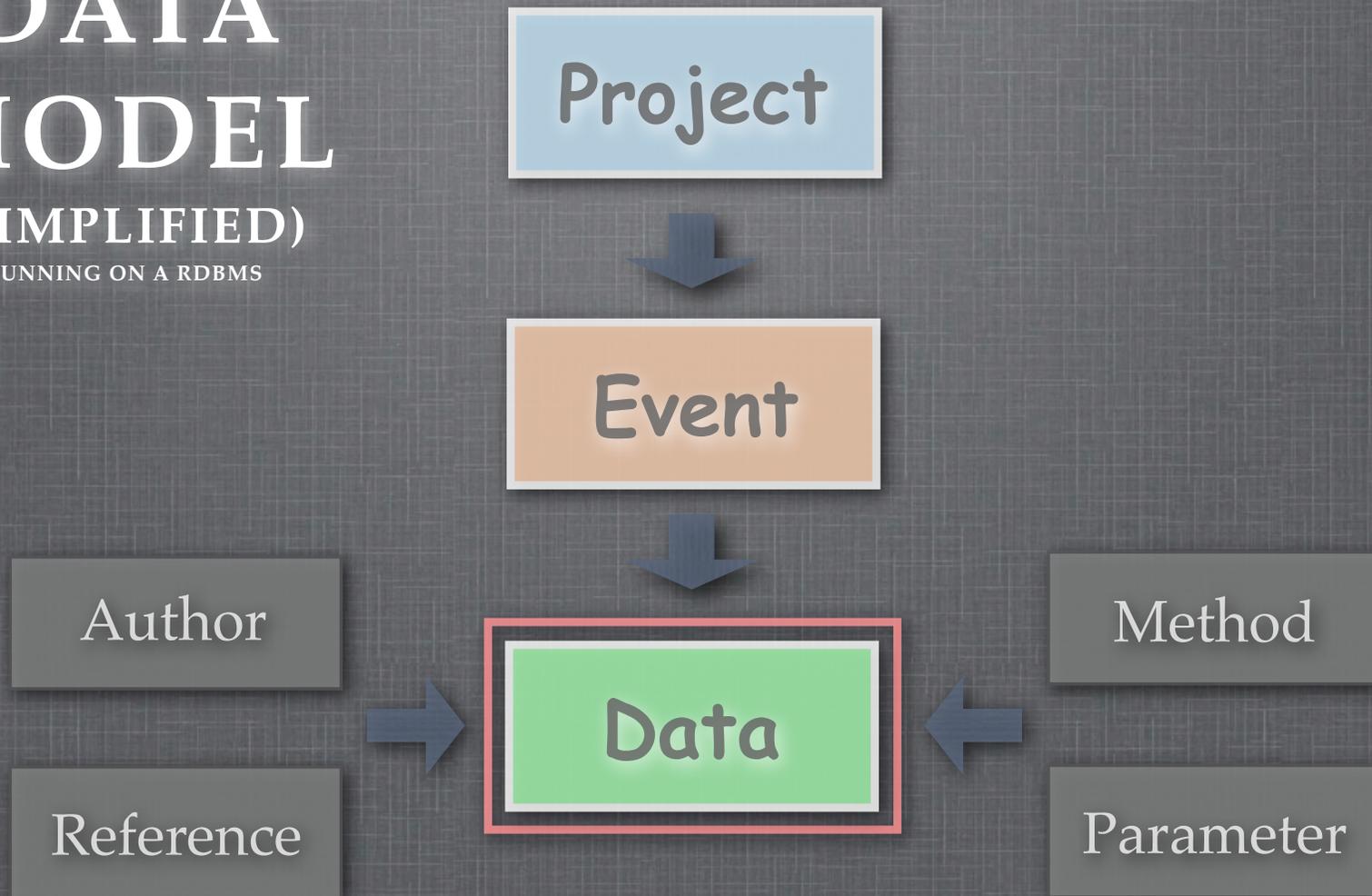
Alfred Wegener Institute for Polar and Marine Research, Bremerhaven
MARUM Center for Marine Environmental Sciences, Bremen

<http://www.wdc-mare.org>



How are metadata and data organized ?

DATA MODEL (SIMPLIFIED) RUNNING ON A RDBMS



DATA CLASSIFICATION

technical	scientific level	<i>example</i>
numbers	primary (raw)	<i>counts</i>
text	secondary	<i>percentage</i>
objects	tertiary	<i>calculations</i> <i>interpretations</i>



GEO-CODE & META-DATA

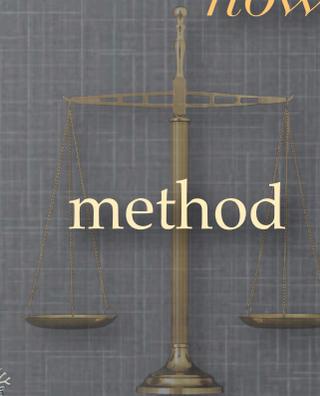
when ?



what ?

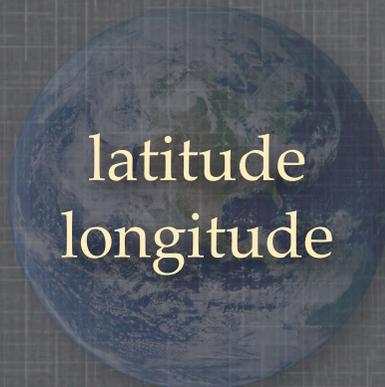
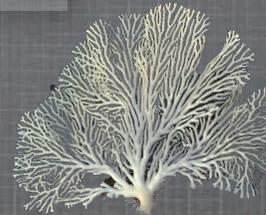


how ?



123.456

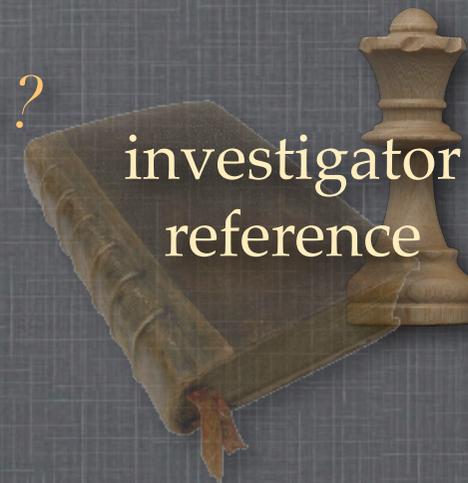
text



where ?

ice, water, air,
sediment, object...

who ?



GEOREFERENCED DATA

Data
table

Value	Latitude	Longitude	3.Spatial dim	Date/Time	Age (kyr)
0.555	-67.6421	18.7659	0.3	2004-03-04	
23.7	-67.6421	18.7659			5.7
123.456	-67.6421	18.7659	112.45		1230
...	...				

Binary
object



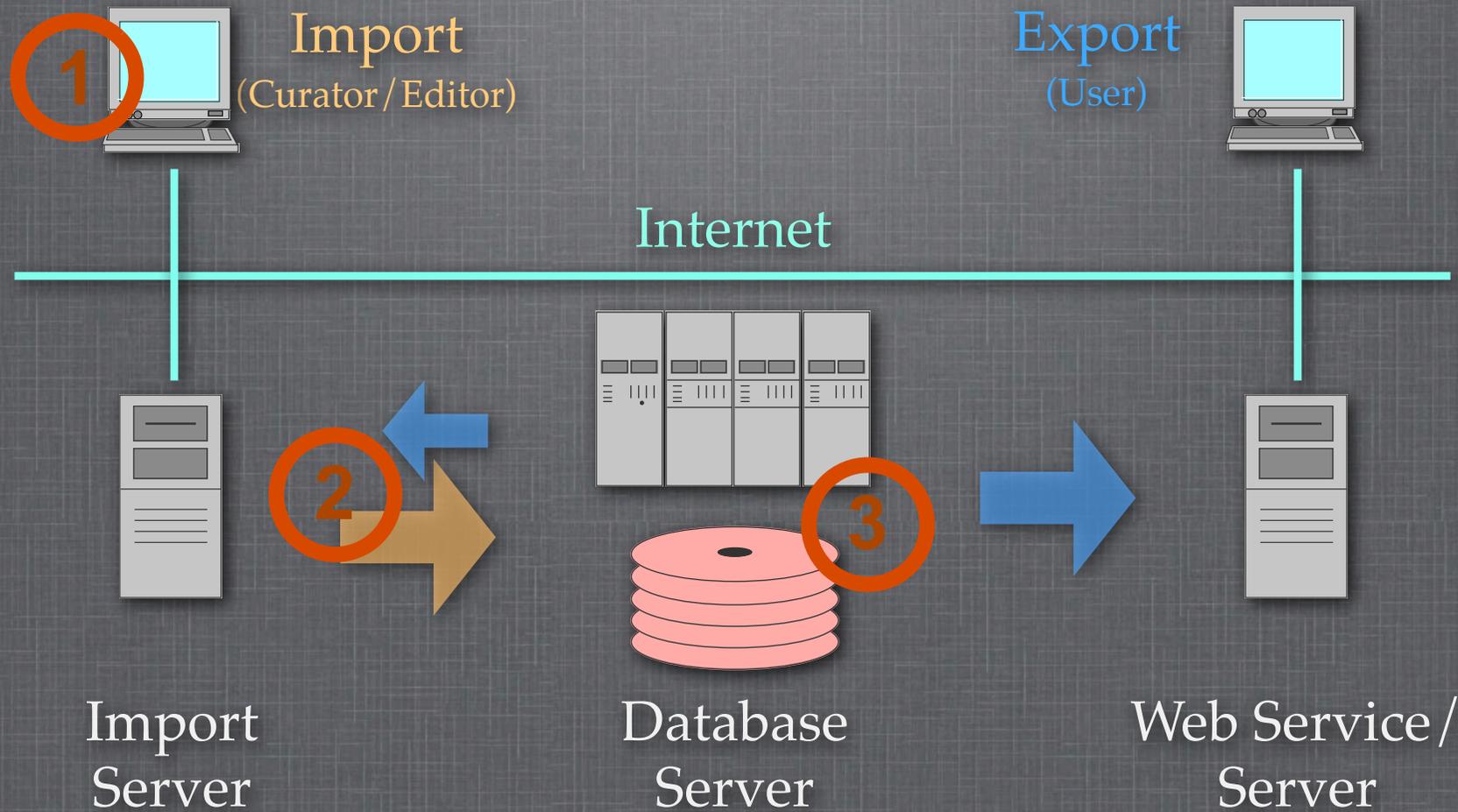
123.456
text
file
...

Latitude	Longitude	3.Spatial dim	Date/Time	Age (kyr)
-67.6421	18.7659	0.3	2004-03-04	



Client-server system

with three-tiered architecture



How can I find and download data ?

Data Access

- 🌐 Pangaea search engine
- 🌐 ART (Advanced Retrieval Tool) > curators only
- 🌐 DDI (Direct Download Interface) > dynamic queries
- 🌐 PanCore > Metadata search
- 🌐 DOI (Digital Object Identifier) > persistent link
- 🌐 Web service > exchange with portals (ISO, DC, XML)
- 🌐 Data-Warehouse > retrieval & compilation



Search engine



Search for:

Search

[Show map](#)

[Help](#)

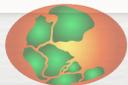
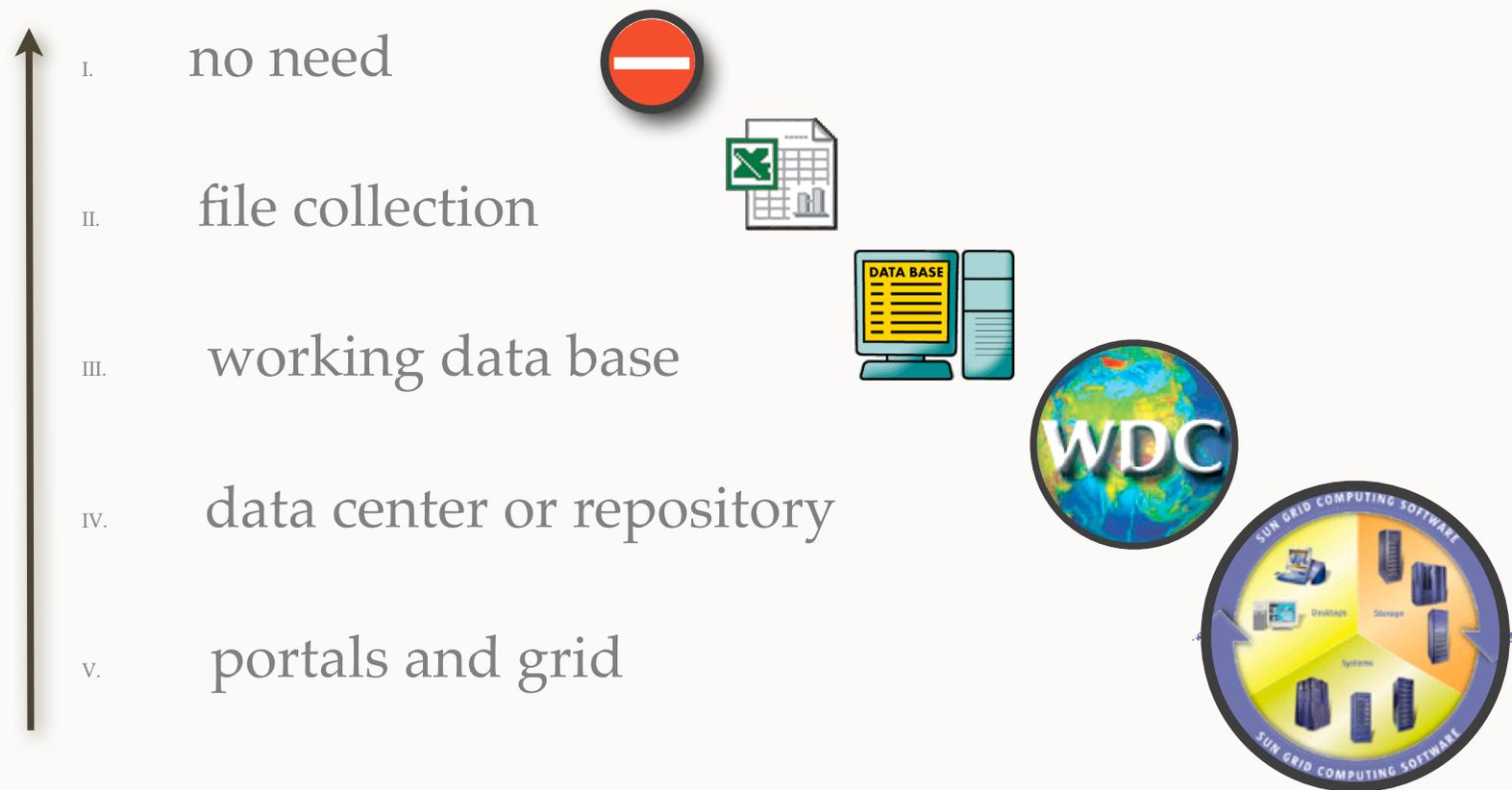
[Info](#)

[Software](#)

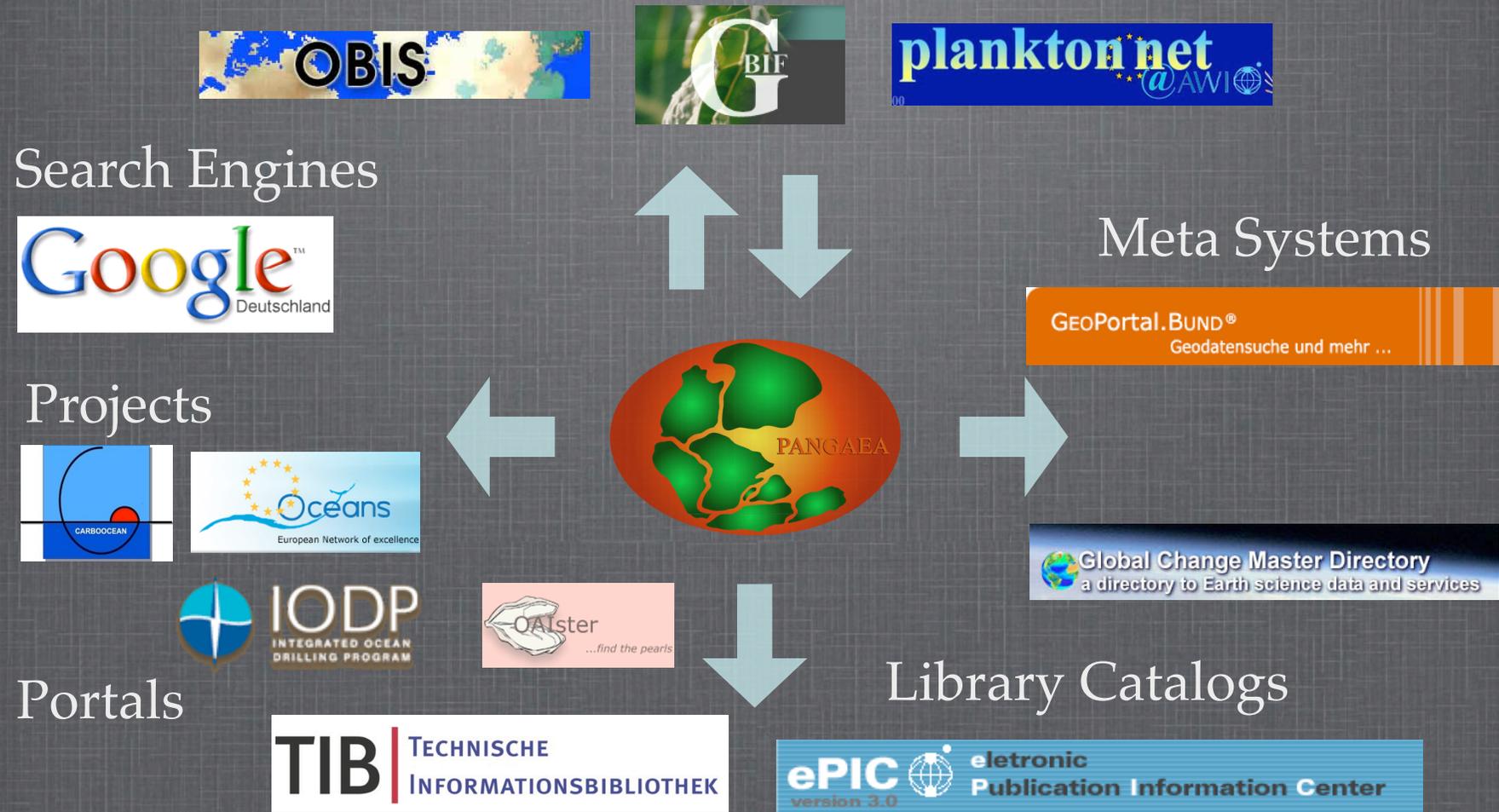
[Advanced search](#)



ENTROPIE IN DATA ARCHIVING



Dissemination of data and metadata via search engines, library catalogs & portals



How can I use data from Pangaea
in my own application ?

DATA MINING

Parameters

- Ammonium [$\mu\text{mol/l}$] (*param52*)
- Nitrate [$\mu\text{mol/l}$] (*param756*)
- Phosphate [$\mu\text{mol/l}$] (*param758*)
- Oxygen [$\mu\text{mol/l}$] (*param754*)
- Silicate [$\mu\text{mol/l}$] (*param755*)

Geographic Coverage

Temporal Coverage

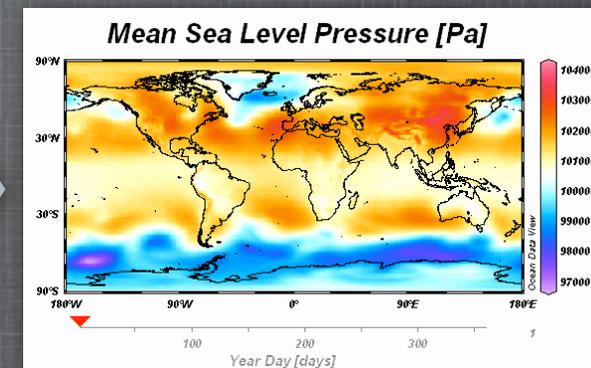
Start date:

End date:

Water Depth

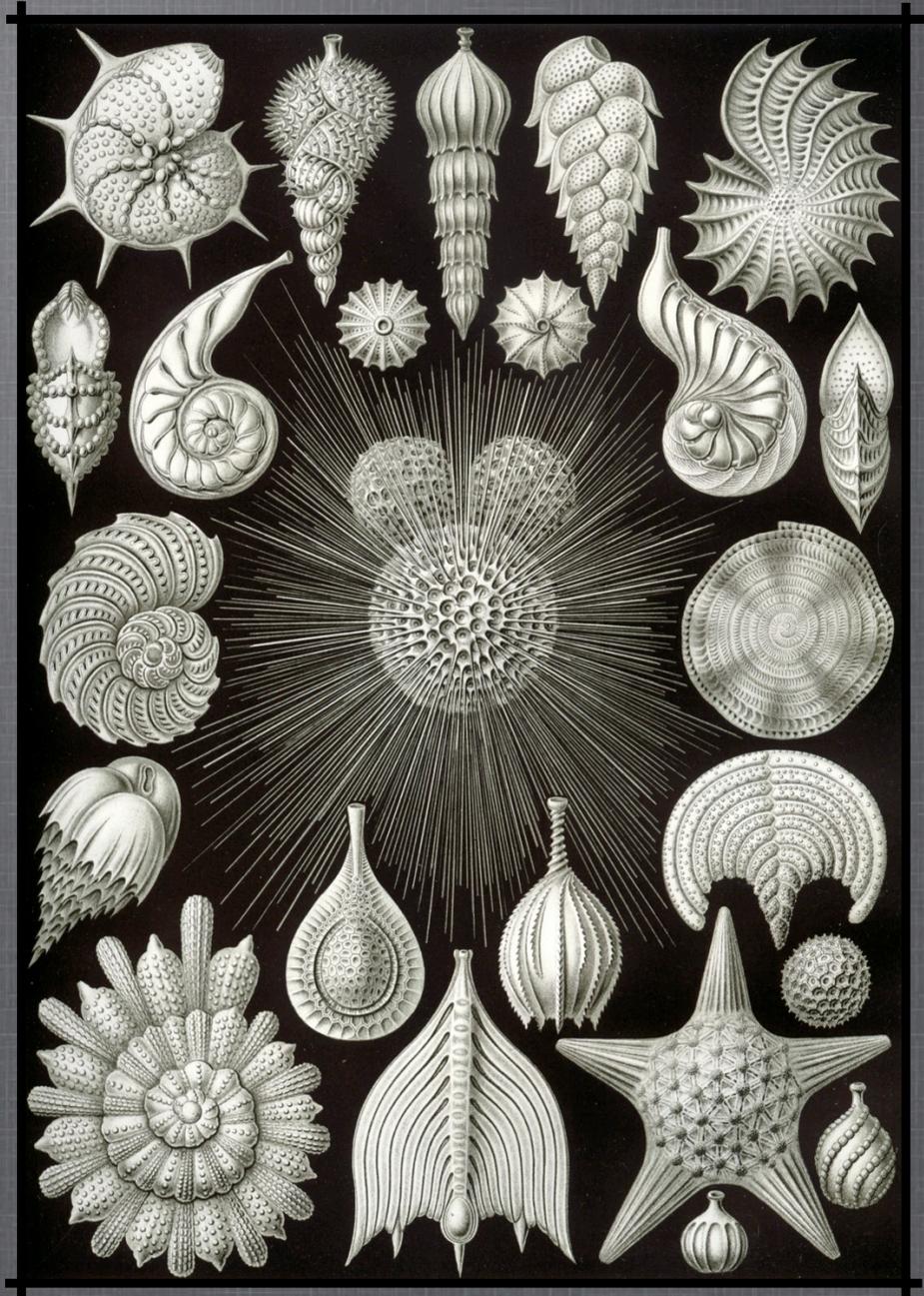
Minimum [m]:

Maximum [m]:

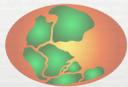
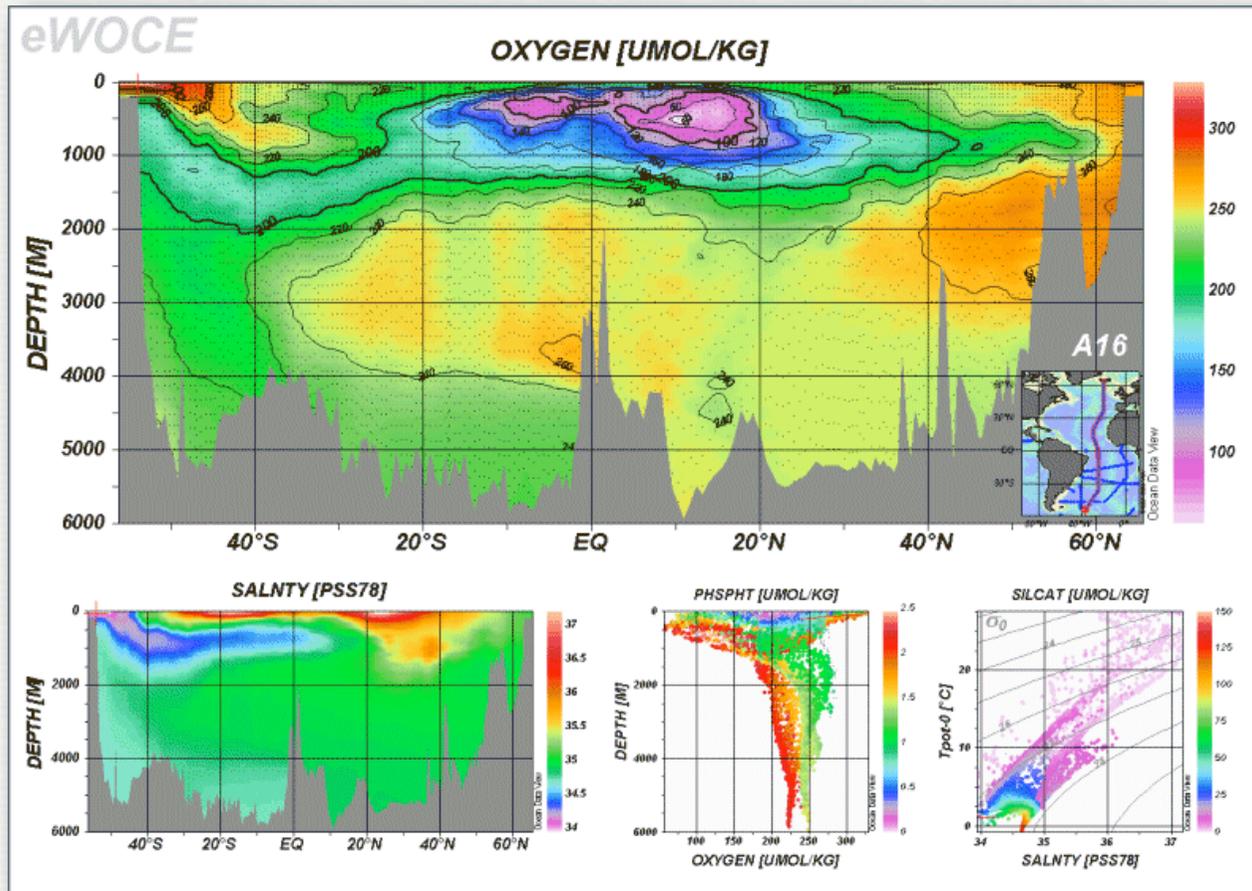


DATA WAREHOUSE

< *Globigerina bulloides* >



Ocean Data View (ODV)



What type of data can I archive / find in
PANGAEA ?

MAJOR PROJECTS

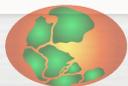
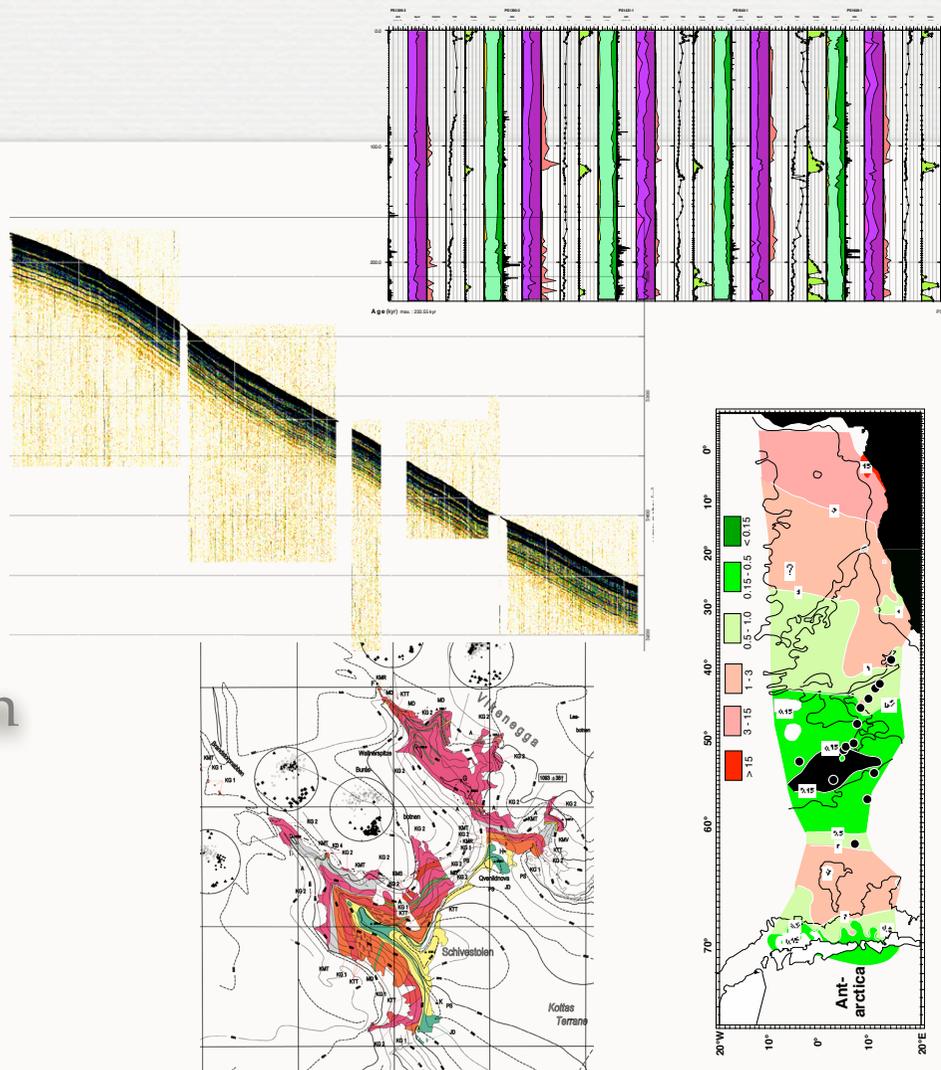
<u>International</u>	<u>EU</u>	<u>National</u>
BSRN	OMARC	rcom
JGOFS	CarboOcean	SIRRO
WOCE	EurOceans	HISTRA
EPICA	HERMES	ARCOD
IODP	EPOCA	DFG/BMBF

<http://www.pangaea.de/projects/>



Examples from Geoscientific Research

- ◆ Sediment profile
- ◆ Seismic profile
- ◆ Mineral distribution
- ◆ Geological map

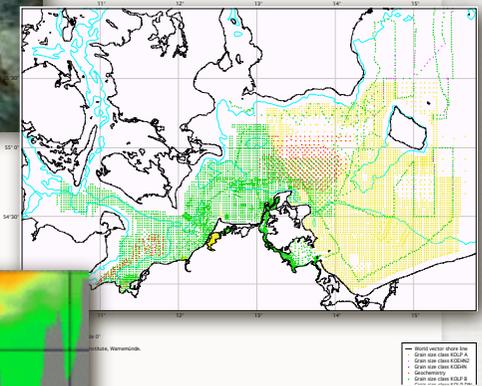


Examples from Environmental Research

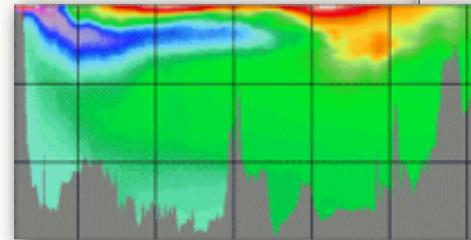
◆ Images



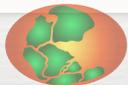
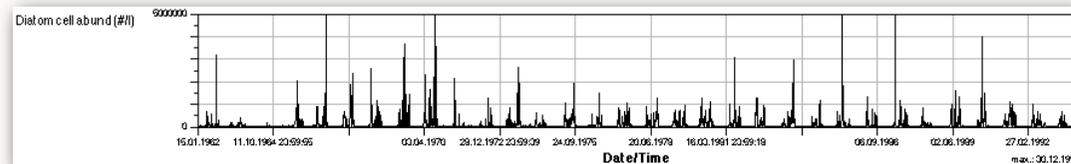
◆ Distributed samples



◆ Oceanographic profiles



◆ Times Series



Air photos

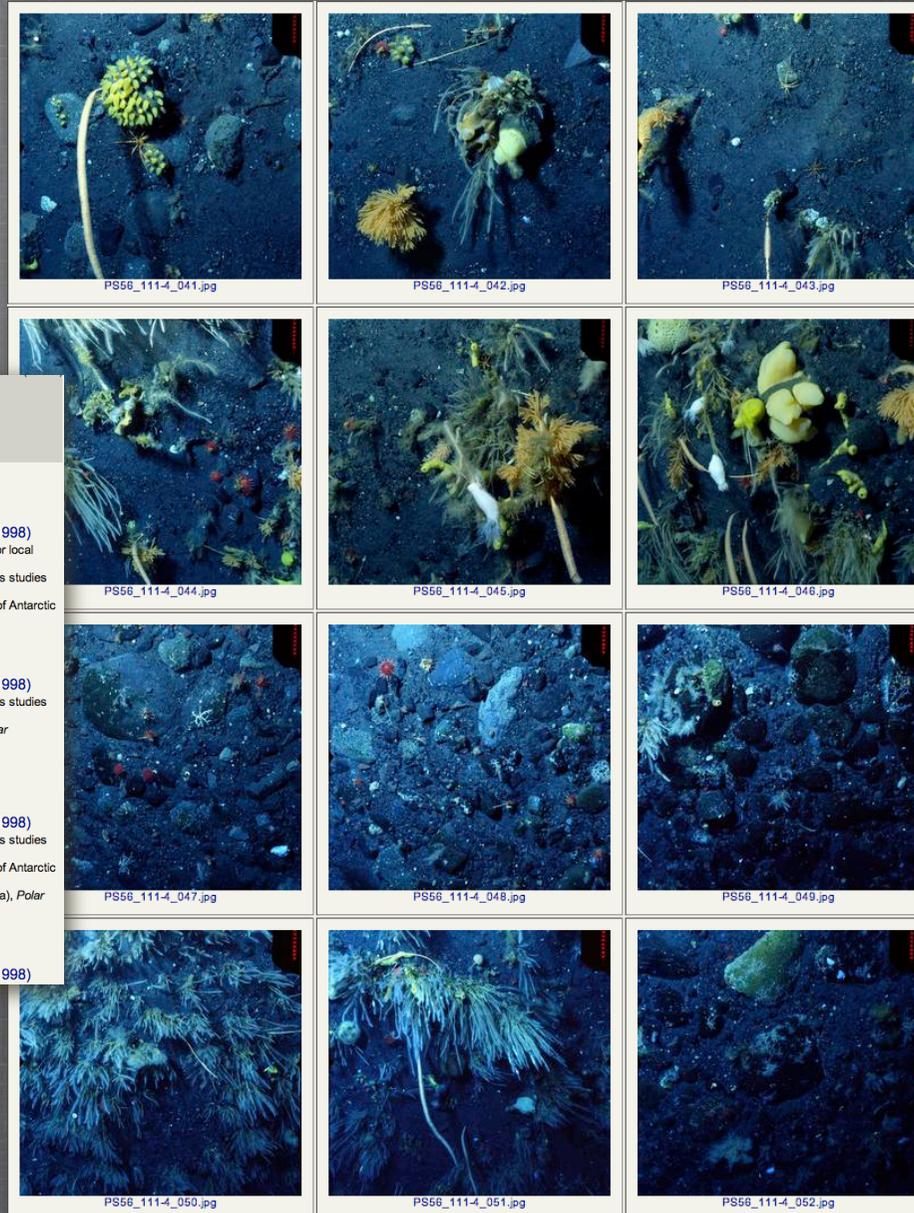


[doi:10.1594/PANGAEA.323540](https://doi.org/10.1594/PANGAEA.323540)



PANGAEA - Publishing Network for Geoscientific & Environmental Data

Sea-bed photos



Search for:
sea-bed
Show map

Search Help Clear

189 datasets found!

<< PREV | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | NEXT >>

- Gutt, J (2004): Sea-bed photographs (benthos) from the Weddell Sea along ROV profile PS48/281 (©AWI, Gutt 1998)**
Reference: Raguà-Gil, JM; Gutt, J, Clarke, A et al. (2004): Antarctic shallow-water mega-epibenthos: shaped by circumpolar dispersion or local conditions? Marine Biology
Gutt, J; Arntz, WE; Balguerías, E et al. (2003): Diverse approaches to questions of diversity: German contributions to benthos studies around South American and Antarctica, Gayana
Gutt, J; Piepenburg, D (2003): Scale-dependent impacts of catastrophic disturbances by grounding icebergs on the diversity of Antarctic benthos, Marine Ecology Progress Series
(and more)
Size: unknown
doi:10.1594/PANGAEA.198686 - Score: 80% - Similar datasets
- Gutt, J (2004): Sea-bed photographs (benthos) from the Weddell Sea along ROV profile PS48/238 (©AWI, Gutt 1998)**
Reference: Gutt, J; Arntz, WE; Balguerías, E et al. (2003): Diverse approaches to questions of diversity: German contributions to benthos studies around South American and Antarctica, Gayana
Gutt, J (2001): High latitude antarctic benthos: a coevolution of nature conservation and ecosystem research?, Ocean and Polar Research
Gutt, J (2001): On the direct impact of ice on marine benthic communities, a review, Polar Biology
(and more)
Size: unknown
doi:10.1594/PANGAEA.198685 - Score: 80% - Similar datasets
- Gutt, J (2004): Sea-bed photographs (benthos) from the Weddell Sea along ROV profile PS48/219 (©AWI, Gutt 1998)**
Reference: Gutt, J; Arntz, WE; Balguerías, E et al. (2003): Diverse approaches to questions of diversity: German contributions to benthos studies around South American and Antarctica, Gayana
Gutt, J; Piepenburg, D (2003): Scale-dependent impacts of catastrophic disturbances by grounding icebergs on the diversity of Antarctic benthos, Marine Ecology Progress Series
Gutt, J; Starmans, A (2001): Quantification of iceberg impact and benthic recolonisation patterns in the Weddell Sea (Antarctica), Polar Biology
(and more)
Size: unknown
doi:10.1594/PANGAEA.198684 - Score: 80% - Similar datasets
- Gutt, J (2004): Sea-bed photographs (benthos) from the Weddell Sea along ROV profile PS48/213 (©AWI, Gutt 1998)**

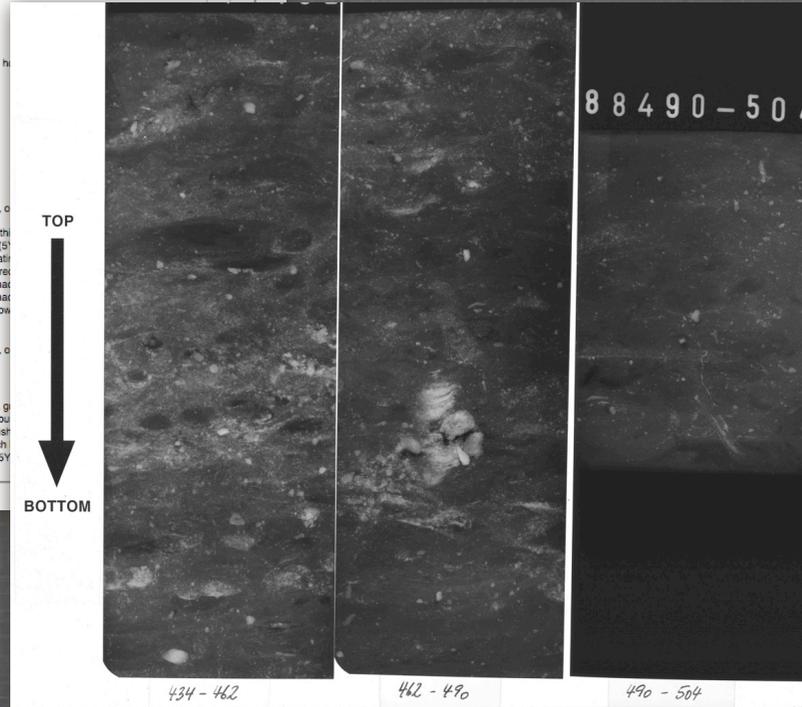
doi:10.1594/PANGAEA.319877



PS1768-8 (SL) North of SW Indian Ridge ANT VIII/3
 Recovery: 8.96 m 52° 35.6' S, 4° 28.5' E Water depth: 3270 m

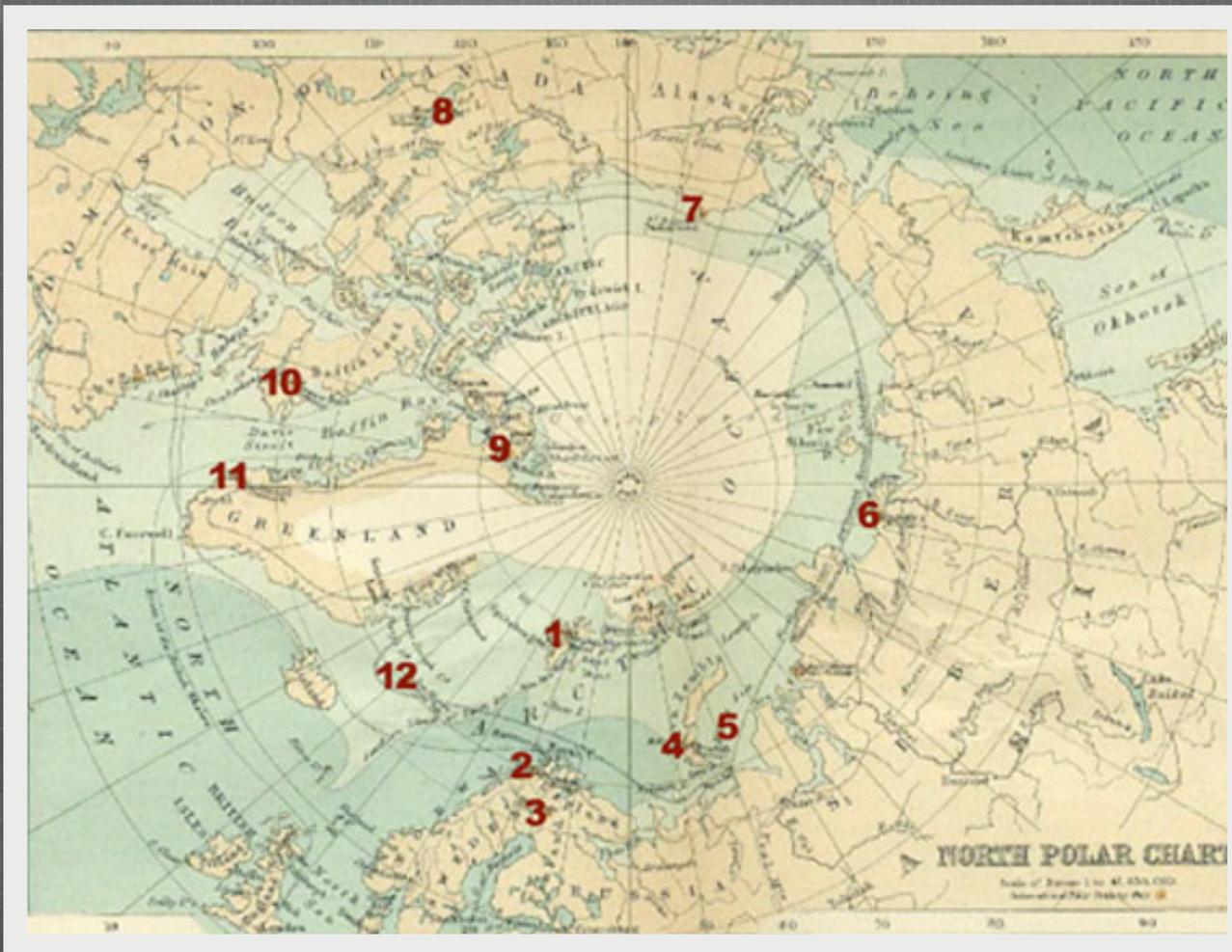
Lithology	Struct. Colour	Description	Age
0	10YR 7/3	0-35 cm: diatomaceous ooze, very pale brown (0-13 cm), light yellowish brown (13-35 cm)	
	10YR 6/4	35-62 cm: diatomaceous ooze, very pale brown (35-53 cm), pale brown (53-62 cm)	
	10YR 7/3	62-70 cm: diatomaceous ooze, very pale brown, two light gray layers (62-64 cm and 66-68 cm)	
	10YR 6/3	70-94 cm: diatomaceous ooze, very pale brown, darker spots	1
1	10YR 7/4	94-139 cm: diatomaceous ooze, light yellowish brown (94-96 cm), dark brown (96-99 cm), pale yellow (99-139 cm)	
	2.5Y 7/4	106-170 cm: partly core deformation	
	5Y 5/3	139-230 cm: diatomaceous mud, homogeneous, olive	
2	5Y 4/2	230-240 cm: diatomaceous mud, h	
3	5Y 5/3	240-440 cm: diatomaceous mud, o occur throughout, 290-306 cm: some th black (5 350-375 cm: alternat scattered diatom 386-387 cm: diatom 395 cm: large burrow	
4	5Y 4/2	440-453 cm: diatomaceous mud, o	
	5Y 4/2	453-486 cm: diatomaceous mud, g 453-458 cm: some bu 474-478 cm: yellowish 480-483 cm: ash-rich 485-486 cm: olive (5Y	
	2.5Y 5/2		
5	5Y 5/3		

Sediment core documentation



doi:10.1594/PANGAEA.108079





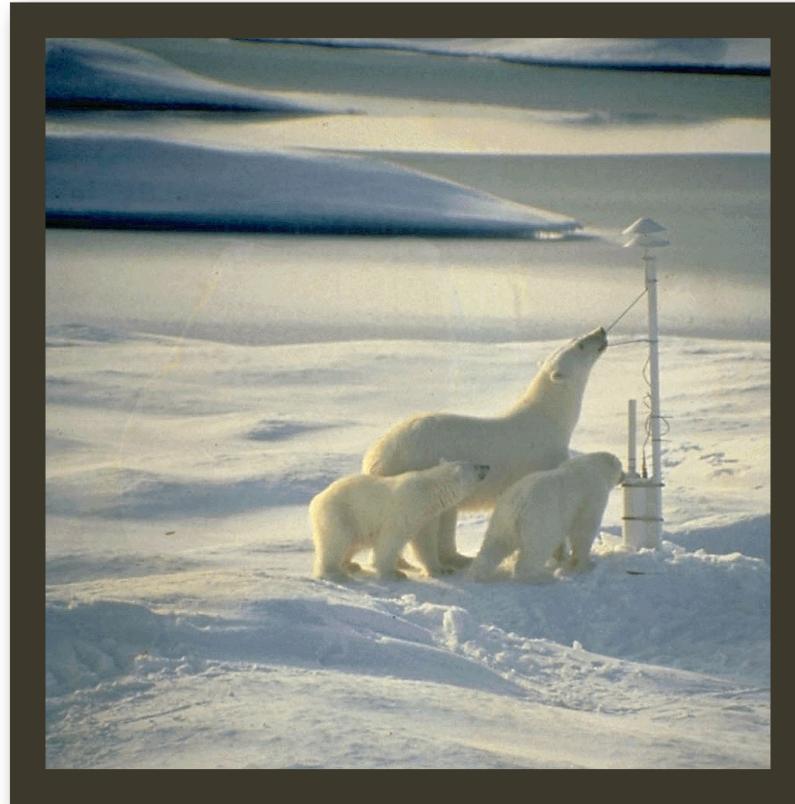
International Polar Year (1882-1883)

A photograph of a rocky coastline. Two prominent rock stacks are visible in the sea. The stack on the left is taller and more rectangular, while the stack on the right is shorter and more jagged. The sky is a clear, pale blue. The foreground shows dark, wet rocks.

Audio record

[doi:10.1594/PANGAEA.339110](https://doi.org/10.1594/PANGAEA.339110)

NO DATA FOUND ?



might not be the problem of the database ...

Contact

info@pangaea.de

This presentation is available at [hdl:10013/epic.31845.d001](https://hdl.handle.net/10013/epic.31845.d001)