



# Forschungsdaten 2010

## Relevanz, Positionen und Akteure

**Sünje Dallmeier-Tiessen**

Vorher

Alfred-Wegener-Institut für Polar- and Meeresforschung (Bremerhaven), Helmholtz Gemeinschaft, Helmholtz Open Access Projekt

Jetzt

CERN, Scientific Information, Open Access und Humboldt Universität zu Berlin, IBI, Berlin

Kontakt: [suenje.dallmeier-tiessen@cern.ch](mailto:suenje.dallmeier-tiessen@cern.ch)

# Agenda

## Relevanz von Forschungsdaten

- “viel Wert” in mehrfacher Hinsicht

## Positionen

- Forschungsförderer, Infrastrukturbetreiber, disziplinspezifische Gremien

## Akteure

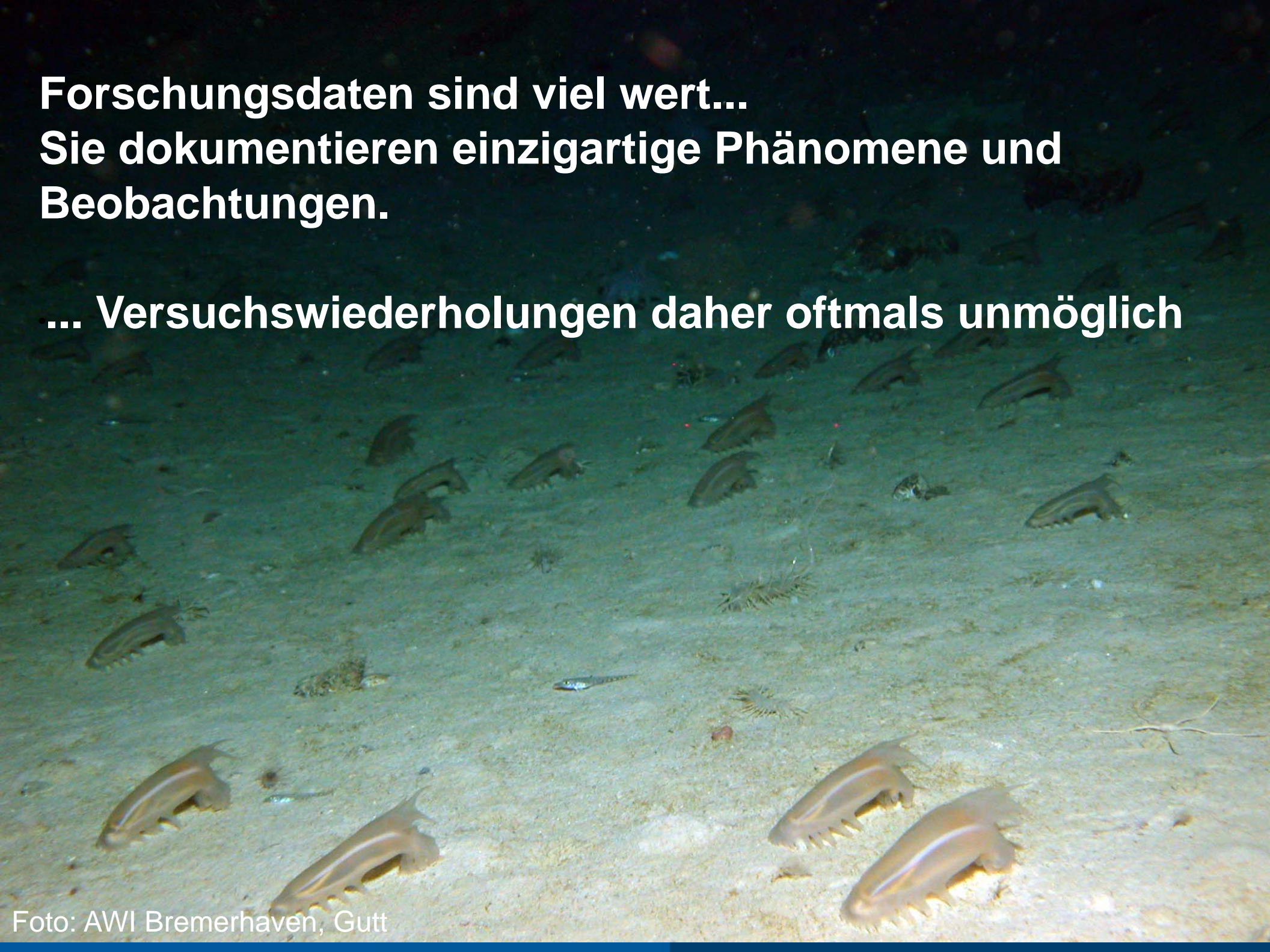
- Wissenschaft, Bibliotheken, Rechenzentren

## Umsetzung

- Rahmenbedingungen
- Modelle – “standalone”, textzentriert, datenzentriert

**Forschungsdaten sind viel wert...  
Sie dokumentieren einzigartige Phänomene und  
Beobachtungen.**

**... Versuchswiederholungen daher oftmals unmöglich**



# Forschungsdaten sind viel wert...

- **Nachnutzbarkeit, Integrität und Nachvollziehbarkeit**
  - Oftmals einmalige Daten
  - Weitere Interpretation der Daten
  - Keine kostenintensiven Messwiederholungen
  - Unabhängige Qualitätskontrolle
- **Forschungsdaten für Lehre und Forschung**
  - In der Schule und im Studium
  - In der Weiterbildung, z.B. von technischem Personal

# Positionen

- Forschungsförderer: z.B.



- Infrastrukturbetreiber

## Auf disziplinspezifischer Ebene

– von Initiativen bis hin zu Policies:

- Unterschiedliche Akteure: Wissenschaftler, Editoren, Infrastruktureinrichtungen
- Von weichen Empfehlungen bis “harten” Bedingungen
- Teilweise mit weiterführenden Empfehlungen zur Umsetzung

z.B. Bermuda Statement – Genomdaten

z.B. DPHEP - Hochenergiephysik

# Akteure

... kennen wir z.T. bereits von den Positionen: Forschungsförderer, Gremien, Infrastrukturbetreiber etc.

## Hands-on:

Bibliothek

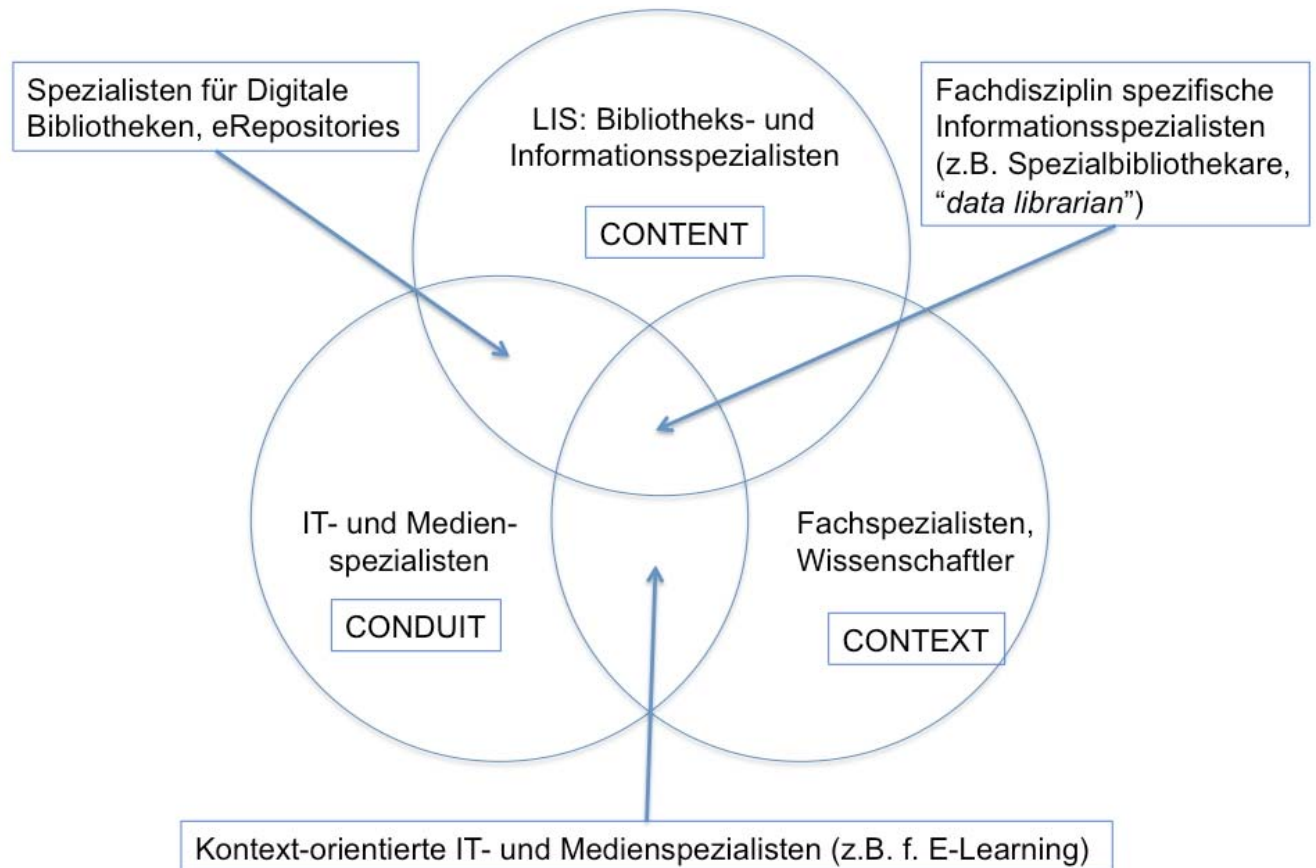
Rechenzentrum

Wissenschaft

Kompetenzen

identifizieren

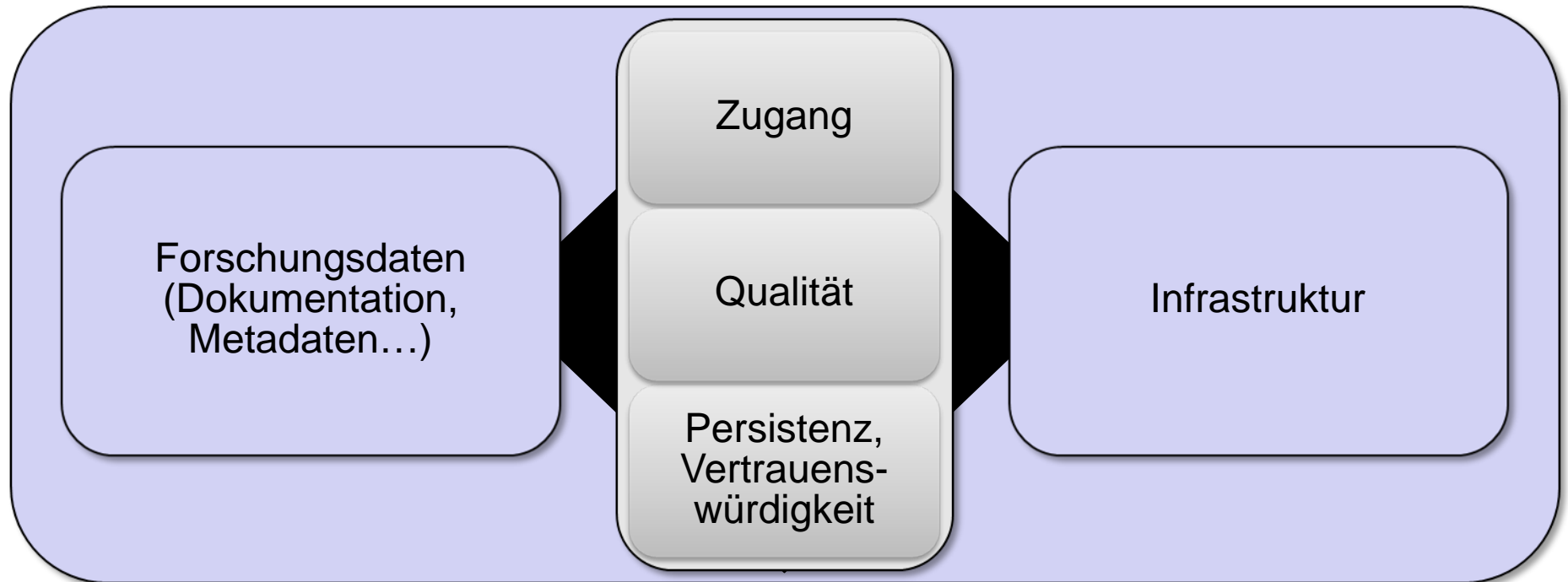
und kombinieren!



Pampel et al., 2009

# Voraussetzung zur Publikation

## Zusammenarbeit von Wissenschaft und Dienstleistern



A  
K  
T  
E  
U  
R  
E

### Wissenschaft:

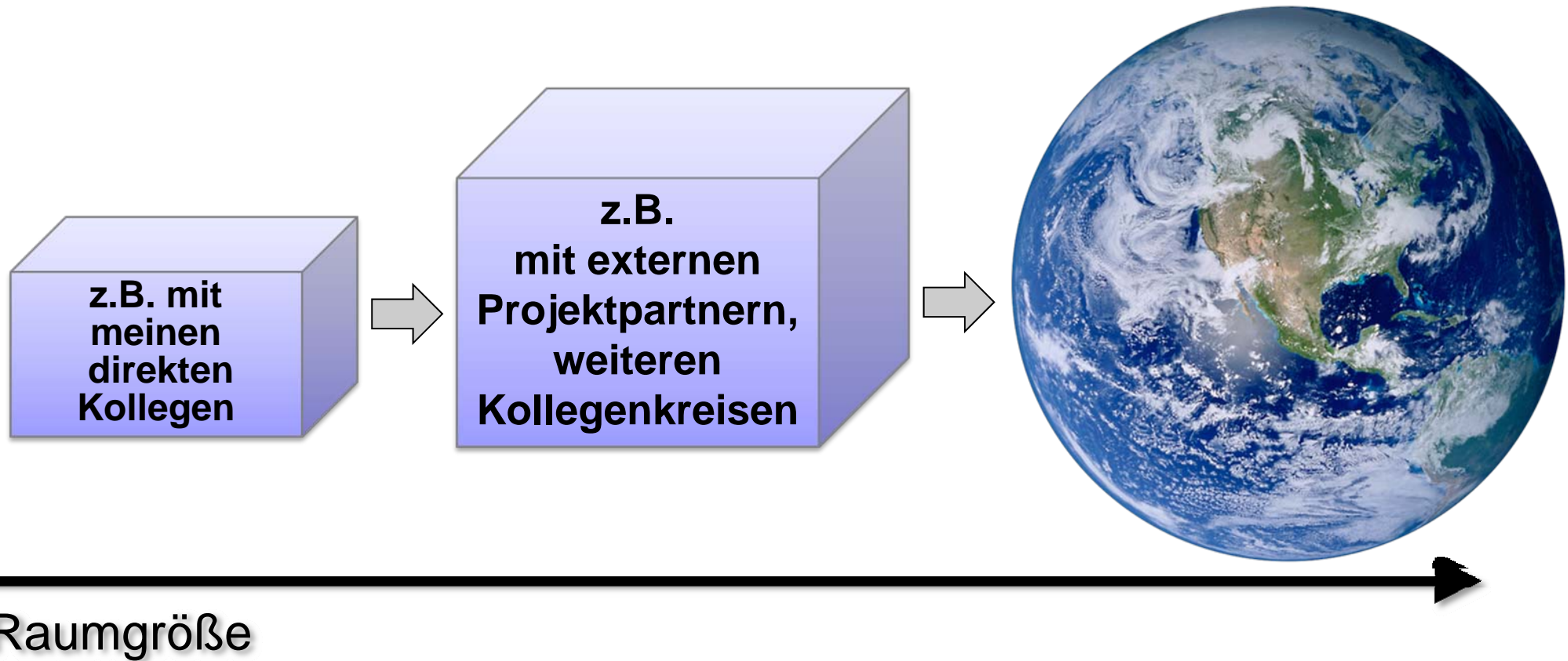
Datenbereitstellung,  
Dokumentation,  
Qualitätskontrolle



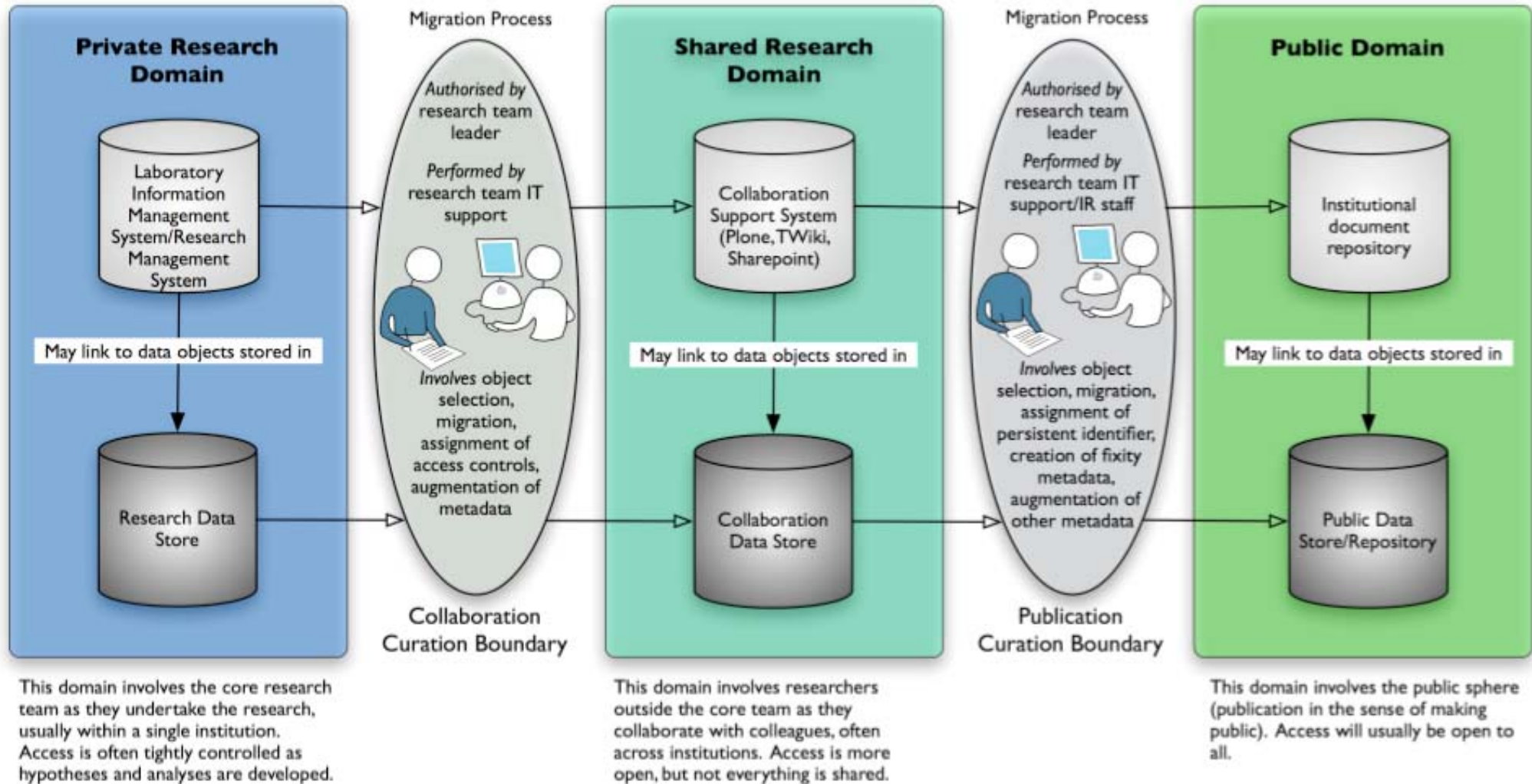
### Infrastruktureinrichtungen:

Infrastrukturangebote,  
Standardisierung,  
Vertrauenswürdigkeit,

# Raum der Datenpublikation



# Raum der Datenpublikation

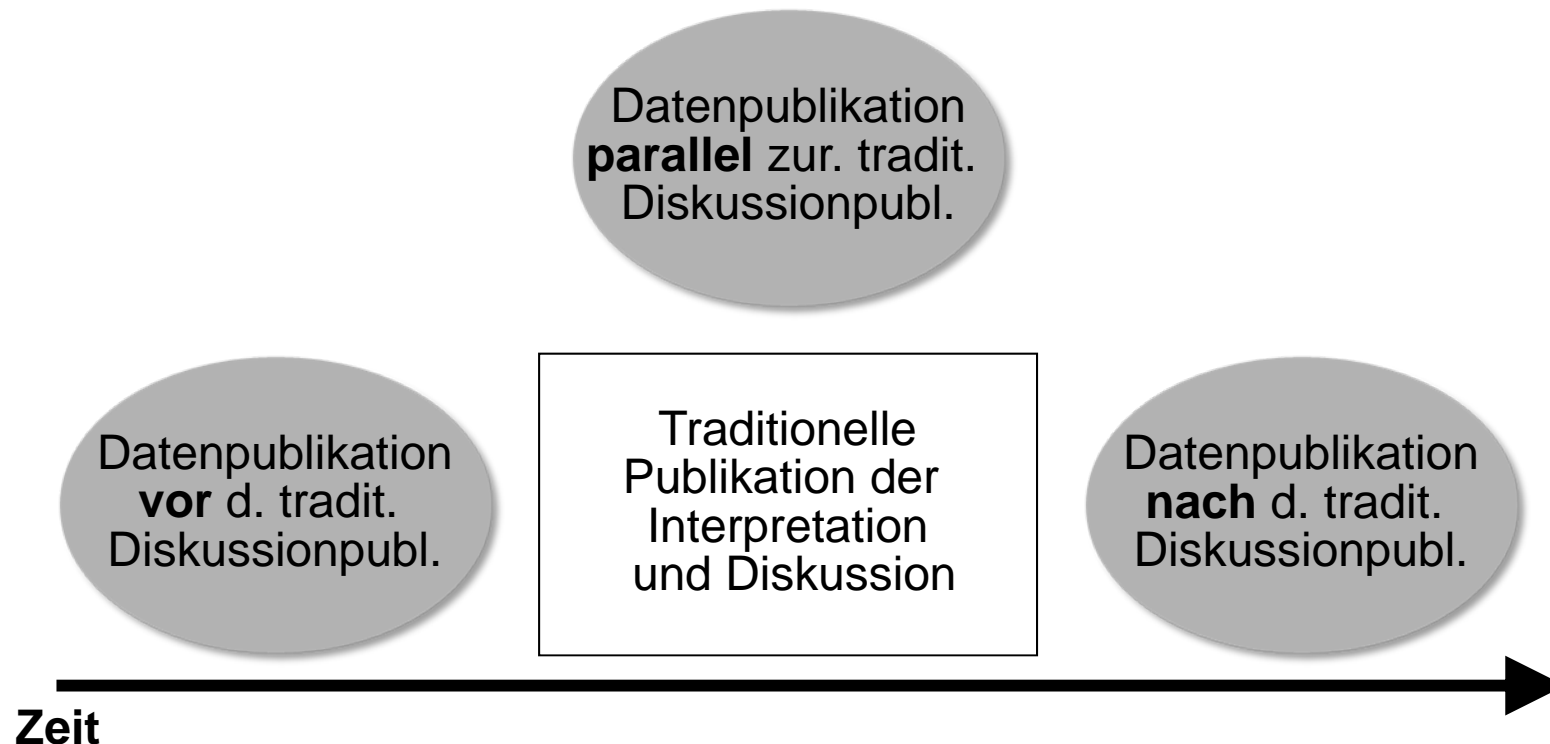


Version 1.4, <http://andrew.treloar.net/>, 07Dec07

# Zeitpunkt der Datenpublikation

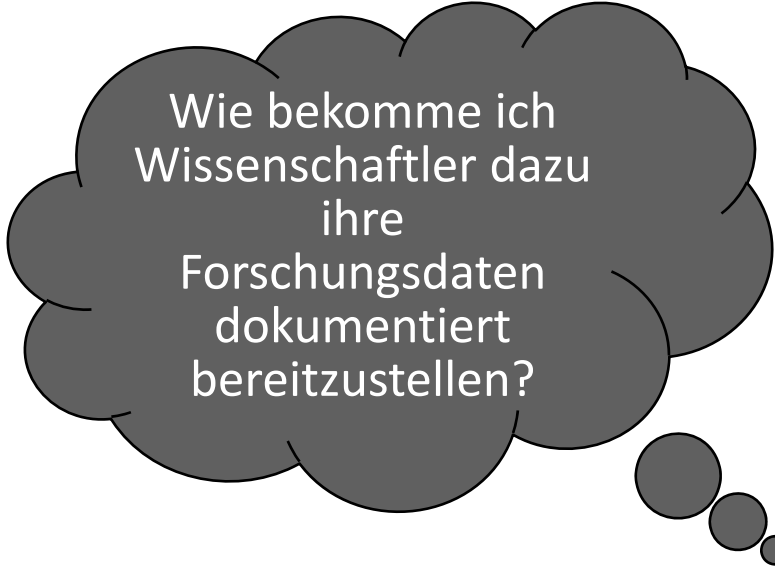
- Direkt nach Datenproduktion/Vor der Interpretation (Pre-)
- Mit der Analyse/Interpretation (with publication)
- Nach der Publikation der Interpretation (Post-)

→ **vielfältige Gründe für die Wahl des Zeitpunktes**

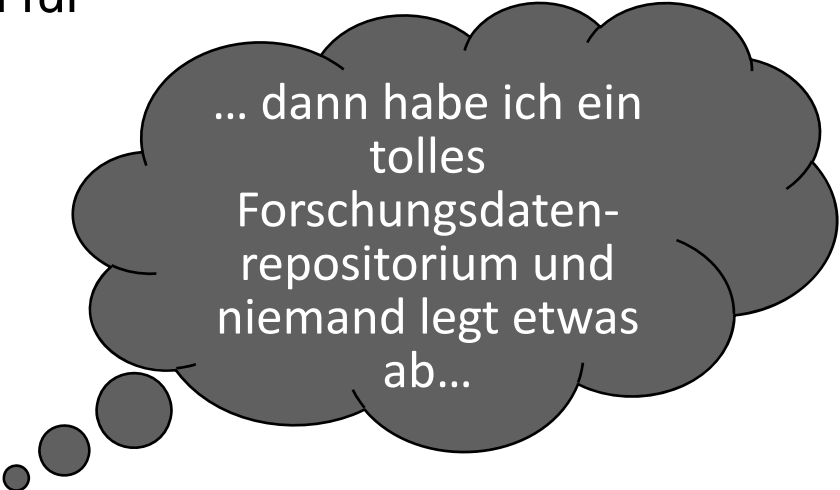


# Was hindert Wissenschaftler an der Publikation von Forschungsdaten – am “Data Sharing”?

- Rechtliche Vorgaben
- Fehlende Anreize für mehr Aufwand
  - Z.B. Wertung als Publikation
- Angst vor “Misuse” and “Misinterpretation”
- Fehlende “eingängige” Zitiermöglichkeiten für Forschungsdaten
- Tlw. fehlende (vertrauenswürdige) Infrastrukturen für die gewünschte Publikationsform
- Angst vor fehlender Kontrolle über die Forschungsdaten
- Unwissenheit: Standardisierung, Workflows etc...
- Fehlende Kooperationen zwischen Akteuren



Wie bekomme ich  
Wissenschaftler dazu  
ihre  
Forschungsdaten  
dokumentiert  
bereitzustellen?



... dann habe ich ein  
tolles  
Forschungsdaten-  
repositorium und  
niemand legt etwas  
ab...

# Umsetzung



## Bereitstellung von Forschungsdaten in der Praxis

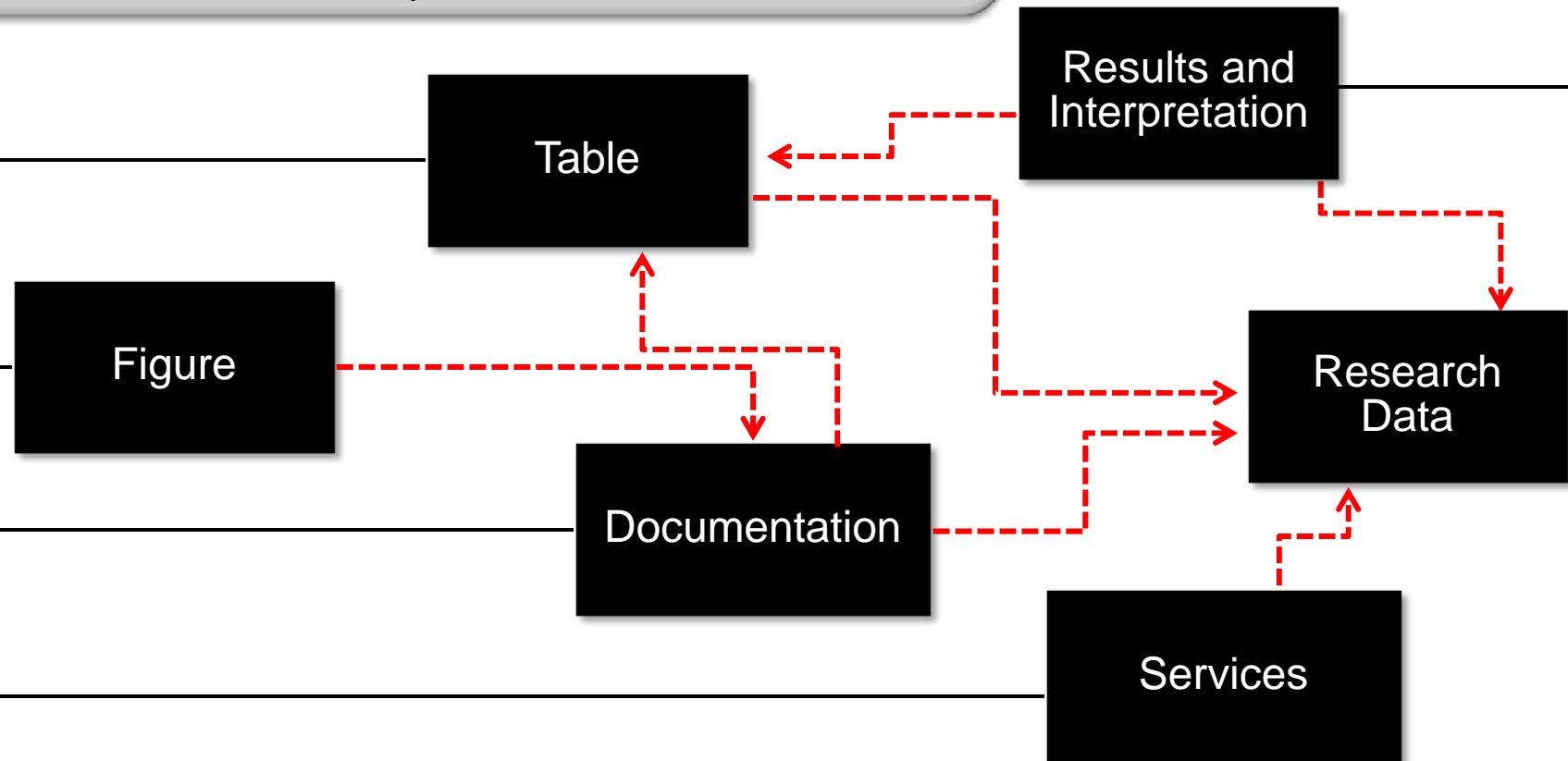
- objektzentriertes
- textzentriertes
- datenzentriertes

### Modell

# Publikationsmodell


## – objektzentriert (“stand alone” & “mash up”)

- je nach Modell oder Kombination variieren die Verantwortlichkeiten der Akteure und der Publikationszeitpunkt
- Wichtig: Persistenz der Komponenten und Modelle



# Weltdatenzentrum: Pangaea

- Publishing Network for Geoscientific and Environmental Data



**PANGAEA®**  
Publishing Network for Geoscientific & Environmental Data

You are not logged in ([LOG IN](#))

Always quote citation when using data!

---

**Data Description**

**Citation:** Spieß, V; Grobe, H (1996): Paleomagnetic measurements on sediment core PS1387-3, doi:10.1594/PANGAEA.51316,  
*In Supplement to: Grobe, Hannes; Mackensen, Andreas; Hubberten, Hans-Wolfgang; Spieß, Volkhard; Fütterer, Dieter K (1990): Stable isotope record and late quaternary sedimentation rates at the Antarctic continental margin, In: Bleil, U & Thiede, J (eds.), Geological History of the Polar Oceans - Arctic versus Antarctic, NATO ASI Series, Kluwer Academic Publishers, Dordrecht, Boston, London, 539-571, Ind. 10013/epic.11660.0001*

**Project(s):** **Paleoenvironmental Reconstructions from Marine Sediments @ AWI** (AWI\_Paleo)

**Coverage:** West: -5.8667 \* East: -5.8667 \* South: -68.7333 \* North: -68.7333  
 Date/Time Start: 1985-12-28T15:48:00 \* Date/Time End: 1985-12-28T15:48:00  
 Minimum DEPTH, sediment: 0.1 m \* Maximum DEPTH, sediment: 10.0 m

**Event(s):** **PS1387-3** (PS08/365) \* Latitude: -68.7333 \* Longitude: -5.8667 \* Elevation: -2416.0 m \* Date/Time: 1985-12-28T15:48:00 \* Recovery: 10.0 m \* Penetration: 11.2 m \*  
 Location: Atka Bay \* Campaign: ANT-IV/3 (PS08) \* Basis: Polarstern \* Device: Gravity corer (Kiel type) \* Comment: upper 30 cm lost in weight, parallel station PS1506

**Parameter(s):**

#	Name	Short Name	Unit	Principal Investigator	Method	Comment
1	DEPTH, sediment	Depth	m			Geocode
2	Susceptibility	Susceptibility	sensor units	Spieß, Volkhard	Susceptibility unit AWI, MS2C, 145 mm	

**Size:** 100 data points

**Data**

Download dataset as tab-delimited text (use the following character encoding: ISO-8859-1: ISO Western (PANGAEA default))

Daten  
zitierbar

1	2
Depth [m]	Susceptibility [sensor units]
0.06000	0.83
0.15000	1.04
0.25000	1.00
0.35000	1.00
0.45000	1.28
0.55000	0.90
0.65000	0.89
0.75000	0.86
0.85000	0.78

# Weltdatenzentrum: Pangaea



**PANGAEA®**  
Publishing Network for Geoscientific & Environmental Data

You are not logged in ([LOG IN](#))

Always quote citation when using data!

## Data Description

[RIS](#) [BIBTeX](#)

*Citation:* Sachs, O et al. (2009): Benthic organic carbon flux and oxygen penetration depth in the Souther Ocean.  
doi:10.1594/PANGAEA.663056,

*Supplement to:* **Sachs, Oliver; Sauter, Eberhard J; Schlüter, Michael; Rutgers van der Loeff, Michiel M; Jerosch, Kerstin; Holby, Ola (2009):** Benthic organic carbon flux and oxygen penetration reflect different plankton provinces in the Southern Ocean. *Deep Sea Research I*, doi:10.1016/j.dsr.2009.02.003

*Parameter(s):*

#	Name	Short Name	Unit	Principal Investigator	Method	Comment
1	Event label	Event				Metadata
2	<a href="#">LATITUDE</a>	Latitude				<a href="#">Geocode</a>
3	<a href="#">LONGITUDE</a>	Longitude				<a href="#">Geocode</a>
4	<a href="#">DATE/TIME</a>	Date/Time				<a href="#">Geocode</a>
5	DEPTH, sediment	Depth				<a href="#">Geocode</a>
6	Depth, bathymetric	Bathy depth				
7	Gear	Gear				
8	Sample Method	Sample Method				
9	Season	Season				

STD-Projekt, DataCite Initiative

# “Eigenständige” Abbildungen

. PlosOne

prior studies. Our results provide novel insights into AML pathogenesis with potential diagnostic, prognostic, and therapeutic implications.



**Figure 1. Tag-based classification method flowchart.**

doi:10.1371/journal.pone.0009466.g001

## RESULTS [Top](#)

### Categorization of Differentially Expressed Genes

A total of 15,809 expression features were available from 25 studies, utilizing 10 different microarray platforms, and comprising a total of 2,744 patient samples ([Table 1](#)). Of the 15,809 expression features, 7,416 were classified as up-regulated, 6,419 were classified as down-regulated, and 1,974 were not classified with respect to an expression direction. A total of 14,385 (91%) expression features could be mapped to a gene symbol in the UCSC hg18 database, which comprised a total of 4,918 genes.

**Table 1. Acute Myelogenous Leukemia expression profiling studies included in analysis.**

doi:10.1371/journal.pone.0009466.t001

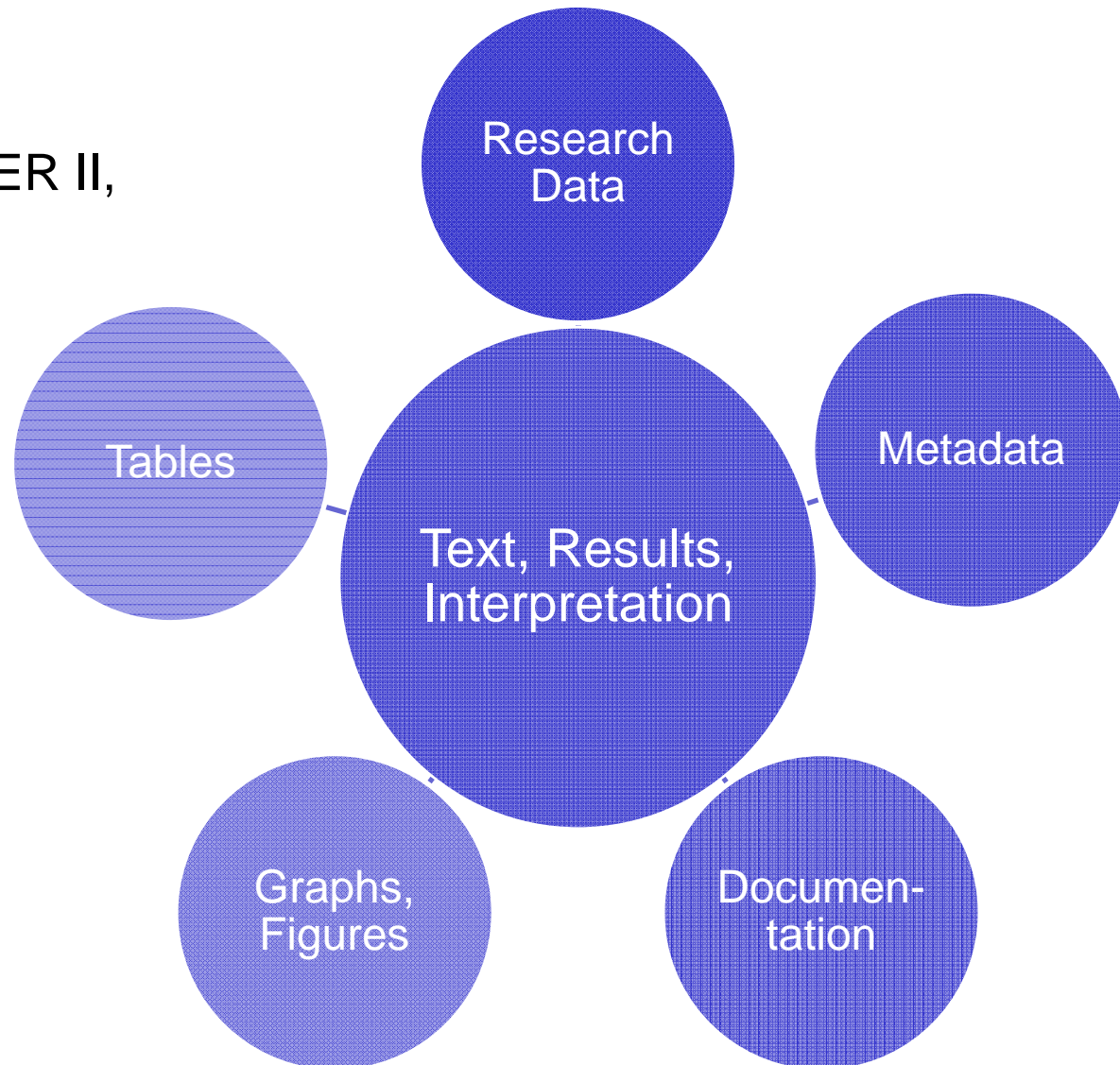
### Standardized Annotation of Gene Expression Features

We annotated each expression feature with standardized identification tags and comparison

doi:10.1371/journal.pone.0009466.g001

# Publikationsmodell - textzentriert

- „traditional“ supplements
- e.g. Enhanced Publications (DRIVER II, SURF):  
“publication of articles along with supplementary data, including the underlying research data, visualisations, public reviews, simulations, and multimedia files.”



# Supplement (files, links...)

 [Publications A-Z index](#) [Browse by subject](#)

**Subscribe** **Register** **Submit manus**

# nature

International weekly journal of science

Search

[Journal home](#) > [Archive](#) > [Brief Communications](#) > [Full text](#) > [Supplementary information](#)

## Supplementary information

From the following article:

[Language evolution: Semantic combinations in primate calls](#)

Kate Arnold & Klaus Zuberbühler

*Nature* **441**, 303(18 May 2006)

doi:10.1038/441303a

[Download plugins and applications](#)

### Supplementary Methods

This file contains Supplementary Figures S1-S5

 [Download PDF file \(402KB\)](#)

### Audio clip 1

A series of 'pyow' calls: these can function as an alarm in response to a nearby leopard, but are also used in other contexts.

 [Download Audio file \(2MB\)](#)

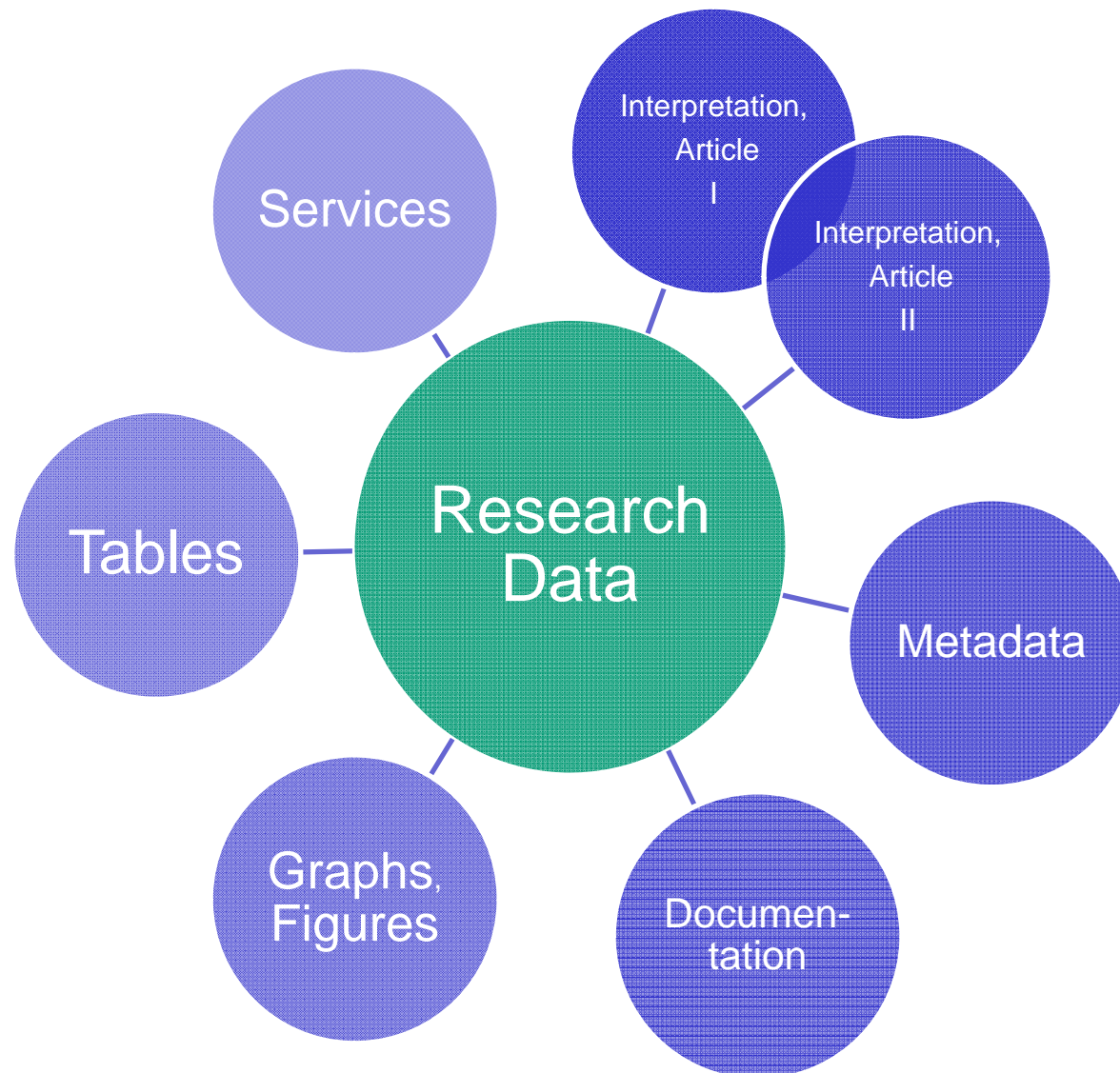
### Audio clip 2

A series of 'hack' calls: mostly functions as an alarm in response to a nearby eagle.

 [Download Audio file \(385KB\)](#)

Quelle: Nature 2006

# Publikationsmodell - datenzentriert



# Datenzentriertes Model - Data Journal

Eigenständige Datenpublikation inkl.  
Dokumentation – qualitätsgesichert durch  
Peer Review und Open Access

Anreiz durch Extra-Publikation

Verlag

Copernicus Publications – OA Publisher



# Data Journal – Repository Reference

## Abstract

On 22 May 1985 the first balloon-borne ozonesonde was successfully launched by the staff of Georg-Forster-Station (70°46' S, 11°41' E). The following weekly ozone soundings mark the beginning of the continuous investigation of Germany to study the vertical ozone distribution in the southern hemisphere.

In 1985 these ozone soundings have been the only record showing the change of vertical ozone distribution in the southern polar stratosphere in September and October. The regular ozone soundings from 1985 until 1992 are a valuable reference data set since the chemical ozone loss became a significant feature in the southern polar stratosphere.

The balloon-borne soundings were performed at the upper air sounding facility of the neighbouring station Novolazarevskaya, just 2 km apart from Georg-Forster-Station. Till 1992, ozone soundings were taken without interruption. Afterwards, the ozone sounding program was moved to Neumayer-Station (70°39' S, 8°15' W) 750 km further west.

## Data coverage and parameter measured

Repository-Reference: doi:10.1594/PANGAEA.547983  
Available at: <http://dx.doi.org/10.1594/PANGAEA.547983>

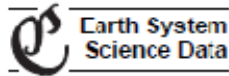
Coverage: East: 11.6300, South: -75.7700  
Location Name: Georg-Forster-Station, Antarctica  
Date/Time Start: 1985-05-22T05:19:00  
Date/Time End: 1992-01-29T01:19:00

Pangaea!

The screenshot shows a blue-themed web interface for the Earth System Science Data (ESSDD) repository. At the top, it displays 'ESSDD' and the issue information '1, 1–13, 2008'. Below this, the title 'Antarctic ozonesonde profiles' and authors 'G. König-Langlo and H. Gernandt' are listed. A navigation menu includes buttons for 'Title Page', 'Abstract', 'Instruments', 'Data Provenance & Structure', 'Tables', and 'Figures'. There are also navigation arrows and buttons for 'Back', 'Close', 'Full Screen / Esc', 'Printer-friendly Version', and 'Interactive Discussion'. At the bottom, a Creative Commons Attribution (CC BY) license logo is visible.

# Data Journal – Akzeptierter Artikel

Earth Syst. Sci. Data, 1, 1–5, 2009  
www.earth-syst-sci-data.net/1/1/2009/  
© Author(s) 2009. This work is distributed under  
the Creative Commons Attribution 3.0 License.



## Compilation of ozonesonde profiles from the Antarctic Georg-Forster-Station from 1985 to 1992

G. König-Langlo and H. Gernandt

Alfred Wegener Institute for Polar and Marine Research, Bussestraße 24, 27570 Bremerhaven, Germany

Received: 20 July 2008 / Published in Earth Syst. Sci. Data Discuss.: 22 September 2008  
Revised: 1 December 2008 – Accepted: 23 December 2008 – Published: 12 January 2009

**Abstract** On 22 May 1985 the first balloon-borne ozonesonde was successfully launched by the staff of Georg-Forster-Station (70°46' S, 11°41' E). The subsequent weekly ozone soundings mark the beginning of a continuous investigation of the vertical ozone distribution in the southern hemisphere by Germany.

The measurements began the year the ozone hole was discovered. They significantly contribute to other measurements made prior to and following 1985 at other stations. The regular ozone soundings from 1985 until 1992 are a valuable reference data set since the chemical ozone loss became a significant feature in the southern polar stratosphere.

The balloon-borne soundings were performed at the upper air sounding facility of the neighbouring station Novolazarevskaya, just 2 km from Georg-Forster-Station. Until 1992, ozone soundings were taken without interruption. Thereafter, the ozone sounding program was moved to Neumayer-Station (70°39' S, 8°15' W) 750 km further west.

### Data coverage and parameter measured

Repository-Reference: doi:10.1594/PANGAEA.547983  
Coverage: East: 11.8300; South: -70.7700;  
Location Name: Georg-Forster-Station, Antarctica  
Date/Time Start: 1985-05-22T05:19:00  
Date/Time End: 1992-01-29T01:19:00

Parameter	Short Name	Unit	Comment
Altitude	Altitude	m	height above mean sea level
Date/Time	Date/Time	universal time code (UTC)	
Longitude	Longitude		at launching point
Latitude	Latitude		at launching point
Ozone, partial pressure	O <sub>3</sub>	mPa	
Pressure, at given altitude	PFPF	hPa	
Temperature, air	TTT	degC	
Wind direction	AD	deg	
Wind speed	W	m/sec	

### 1 Introduction

The first permanently operated German research base – later named Georg-Forster-Station – was established in 1976 in the Schirmacher Oasis at 70°46' S, 11°41' E. The station was permanently used and operated as an annex to the Russian station Novolazarevskaya until 1987, and then as a German Antarctic station named after the German natural scientist, author and revolutionary Georg Forster (1754–1794) until 1993.

Long-term studies of magnetospheric-ionospheric processes, geophysical investigations, biological studies and sea ice observations using satellite imaging were performed.

The station became known to the international scientific community when the vertical extent of the “ozone hole” in the southern polar stratosphere was firstly recorded by regular balloon-borne ozone observations in 1985 (Gernandt, 1987a, b).

The ozone sounding programme was a major contribution of the Meteorological Service to the Antarctic research of the German Democratic Republic (GDR). The station was established as a long-term ozone-sonde observatory in cooperation with the Russian Arctic and Antarctic Research Institute (AARI) and the Aerological Observatory Lindenberg (AOL) in order to study the climatology of the ozone layer in

Datensatz auf bekannte Art und Weise zitierbar

Datensatz nachnutzbar:

Daten qualitätsgesichert und online abrufbar

Publikation des Datensatzes als eigenständiger Artikel

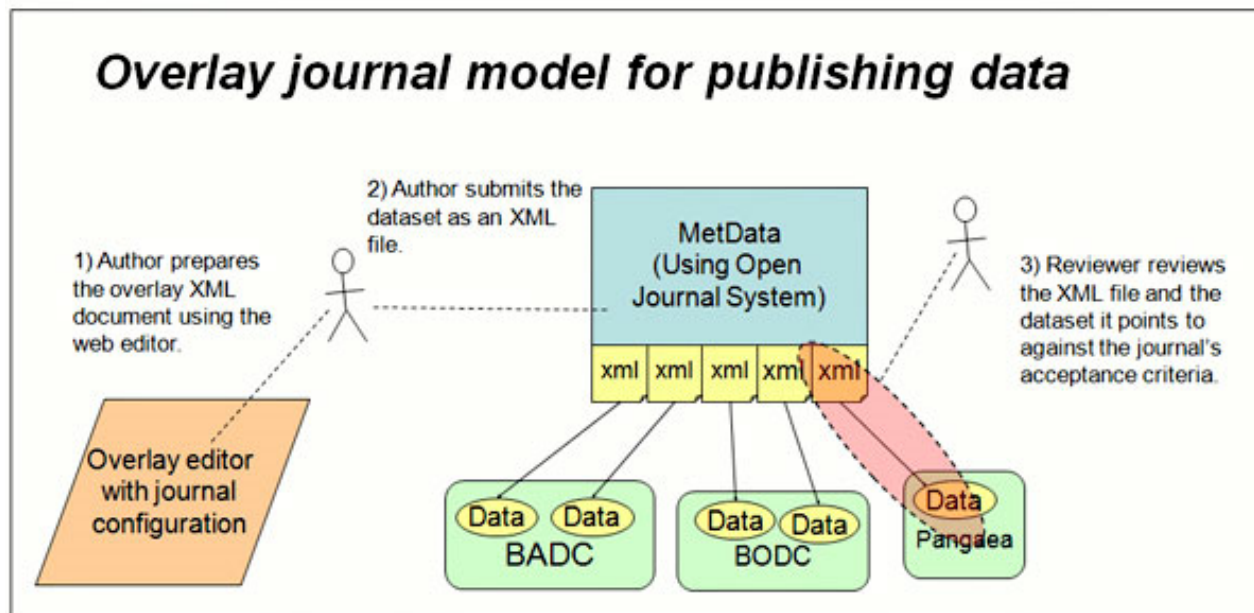
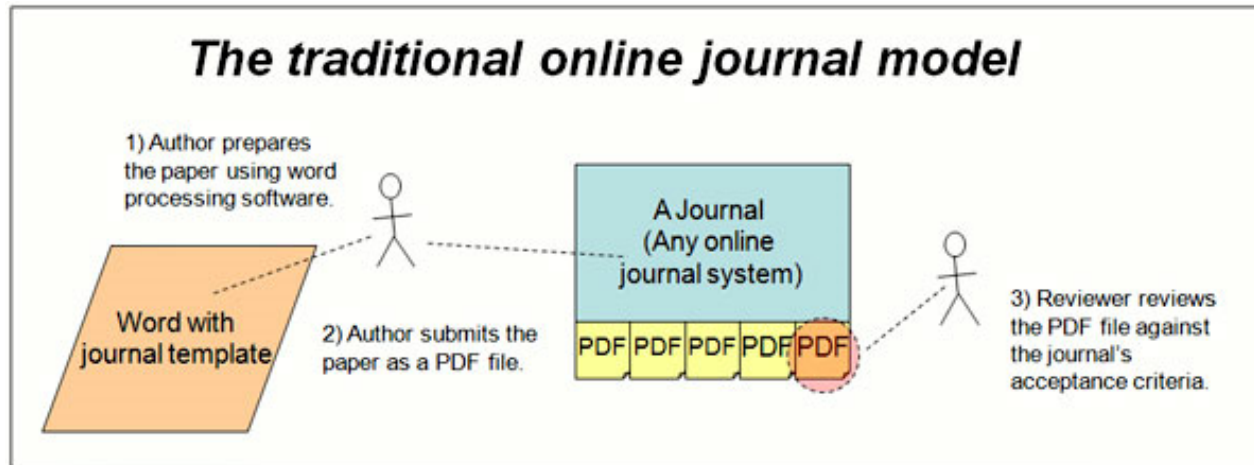


Correspondence to: G. König-Langlo  
([gert.koenig-langlo@awi.de](mailto:gert.koenig-langlo@awi.de))

Published by Copernicus Publications

<http://www.earth-system-science-data.net/>

# Datenzentriertes Modell: - Overlay Data Journal, e.g. OJIMS



**Adding and combining services**

Callaghan et al., 2009

# Zusammenfassung

## - Rolle der Bibliotheken

- Verschiedene Positionen und Initiativen zum Thema Forschungsdaten – gemeinsame und disziplinspezifische Wege
- Bekannte Akteure mit tlw. neuen Aufgaben und neuen Kooperationen

### Praxis:

- Starke disziplinspezifische Anforderungen im Umgang mit Forschungsdaten → bei allen Ideen von Anfang an mit der aktiven Wissenschaft kooperieren
- Modelle für die Publikation der Forschungsdaten,
  - Charakteristika und Flexibilitäten beachten
  - unterschiedliche Auswirkungen auf Raum und Zeitpunkt der Datenpublikation
  - Zugang, Qualität und Persistenz – Aufgaben der Akteure

# Referenzen

- Callaghan et al., 2009: Overlay Journals and Data Publishing in the Meteorological Sciences, *Ariadne*, 60, July 2009
- Earth System Science Data (Journal): <http://www.earth-system-science-data.net/>
- Nature Online: [www.nature.com](http://www.nature.com)
- Pampel, H.; Bertelmann, R.; Hobohm, H.-C.: „Data Librarianship“ – Rollen, Aufgaben, Kompetenzen. Tagungsband des 98. Deutschen Bibliothekartag 2009.
- Pangaea: [www.pangaea.de](http://www.pangaea.de)
- PloSOne: [www.plosone.org/](http://www.plosone.org/)
- SURF: [www.suffoundation.nl](http://www.suffoundation.nl)
- Treloar, A.: Data management and the curation continuum: how the Monash experience is informing repository relationships. Vala Conference 2008  
[http://www.valaconf.org.au/vala2008/papers2008/111\\_Treloar\\_Final.pdf](http://www.valaconf.org.au/vala2008/papers2008/111_Treloar_Final.pdf)

Suenje Dallmeier-Tiessen,

Kontakt: [suenje.dallmeier-tiessen@cern.ch](mailto:suenje.dallmeier-tiessen@cern.ch)

Gefördert durch Wolfgang-Gentner-Stipendium, BMBF