

# Persistent Identifiers in the Publication and Citation of Scientific Data

Jens Klump

Helmholtz Centre Potsdam

German Research Centre for Geosciences (GFZ)

IZA Bonn

ISCD Red Cube Seminar

2009-10-28

# STD-DOI Project Team

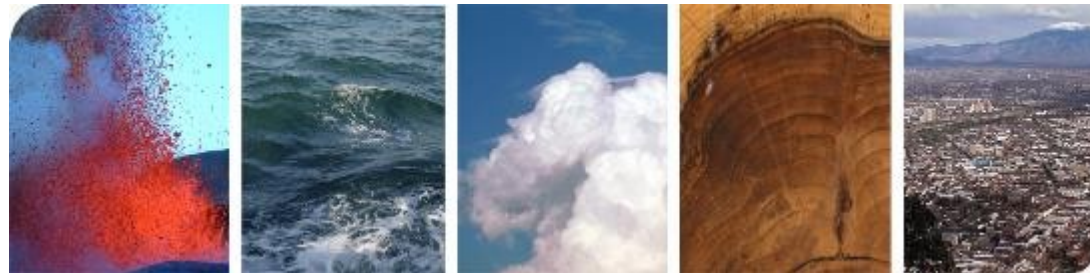
- Jan Brase and Irina Sens, *German National Library for Science and Technology (TIB Hannover)*
- Michael Diepenbroek and Uwe Schindler, *Univ. Bremen, WDC-MARE*
- Hannes Grobe, *AWI Bremerhaven, WDC-MARE*
- Beate Hildenbrand, *DLR-DFD, WDC-RSAT*
- Jens Klump, *GFZ Potsdam*
- Michael Lautenschlager and Heinke Höck, *MPI MET Hamburg, WDC Climate*
- ... and co-workers.

# Agenda

- About GFZ
- Why access to data?
- Why publication of data?
- Why persistent identifiers?
- DFG project STD-DOI
- The practice of data publication
- The future: DataCite

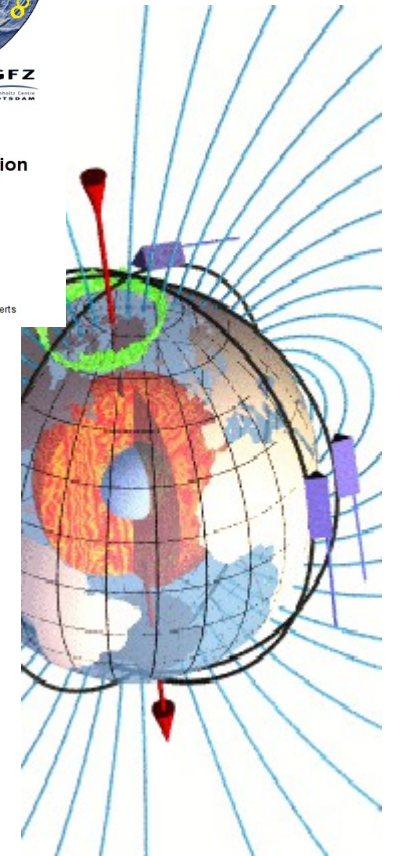
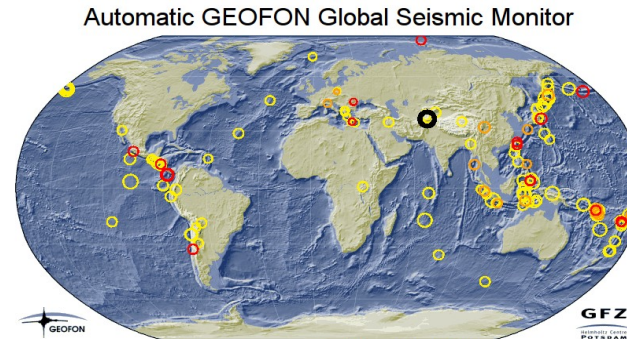
# About GFZ

- Helmholtz Centre Potsdam German Research Centre for Geosciences (GFZ)
- Member of the Helmholtz Association of Research Centres
- Founded in 1992
- Staff: 928 (incl. 413 scientists and 83 PhD students)
- Budget: ca. 76 mio. €



# About GFZ

- GFZ operates analytical infrastructure, drilling equipment, satellites, observatories and other environmental monitoring networks for the scientific community.



# Why Access to Data?

nature

[www.nature.com/nature](http://www.nature.com/nature)

Vol 461 | Issue no. 7261 | 10 September 2009

## Data's shameful neglect

Research cannot flourish if data are not preserved and made accessible. All concerned must act accordingly.

**M**ore and more often these days, a research project's success is measured not just by the publications it produces, but also by the data it makes available to the wider community. Pioneering archives such as GenBank have demonstrated just how powerful such legacy data sets can be for generating new discoveries — especially when data are combined from many laboratories and analysed in ways that the original researchers could not have anticipated.

All but a handful of disciplines still lack the technical, institutional

also the software that will help investigators to do this. One important facet is metadata management software: tools that streamline the tedious process of annotating data with a description of what the bits mean, which instrument collected them, which algorithms have been used to process them and so on — information that is essential if other scientists are to reuse the data effectively.

Also necessary, especially in an era when data can be mixed and combined in unanticipated ways, is software that can keep track of

# The other side ...

The screenshot shows the homepage of stern.de with a search bar and navigation menu. The main article is titled "Norweger hat Krebs-Studie frei erfunden" (Norwegian has invented cancer study). The article text states: "Der norwegische Mediziner und Krebsforscher Jon Sudbø hat die Veröffentlichung gefälschter oder frei erfundener Forschungsergebnisse in drei international angesehenen Fachjournalen gestanden. Wie der Rundfunksender NRK in Oslo am Montag berichtete, reagierte der Wissenschaftler am Radium-Hospital damit auf Vorwürfe zu einem Artikel mit frei erfundenen Forschungsdaten in der britischen Zeitschrift 'The Lancet'. Darin hatte Sudbø die angeblich vorbeugende Wirkung bestimmter schmerzstillender Mittel bei Patienten mit Mundhöhlenkrebs nachweisen wollen. 'Lancet'-Chefredakteur Richard Horton hatte den Fall als 'beispiellosen Schwindel' eingestuft und einen Widerruf veröffentlicht."

Below the article, there is a sidebar with a section titled "MEHR Z" and a sub-section "FÄLSCHUNG VON STUDIEN". The main headline in this section is "Schmerzforscher im Pharnasumpf" (Pain researcher in the swamp). The text continues: "Scott Reuben galt als Pionier der Schmerzforschung. Seinen Empfehlungen folgten Ärzte weltweit: Aber 21 seiner Studien sind frei erfunden". A quote follows: "Das ist ein Skandal ohne Beispiel", sagt Edmund Neugebauer. Der Professor für Chirurgische Forschung an der Universität Witten-Herdecke findet für den einst angesehenen US-Wissenschaftler Scott Reuben deutliche Worte. "Eine so schwere Verfehlung ist nicht wieder gut zu machen. Reuben ist damit wissenschaftlich tot und gehört entlassen." Gemeint ist ein Betrugsfall, der derzeit vor allem die Mediziner in den USA schockiert. Seit 1996 hat der Anästhesist Scott Reuben vom Baystate Medical Center in Springfield, Massachusetts, mindestens 21 seiner Studien schlichtweg frei erfunden. Viele seiner gefälschten Ergebnisse sind in Leitlinien für Ärzte eingeflossen.

Deutsche  
Forschungsgemeinschaft

Sicherung guter  
wissenschaftlicher  
Praxis

Safeguarding Good  
Scientific Practice

Denkschrift



*Eos*, Vol. 90, No. 32, 11 August 2009

# FORUM

## Documenting Precision and Accuracy in the Open Data Policy Era

PAGE 276

How often do research papers discuss measurement uncertainty or present error bars in observational or statistical analysis figures?

the words "error," "uncertainty," "statistical," and "deviation."

Fifty-five percent of the articles did not mention measurement uncertainty or present error bars in data figures. Of the 47 articles that did mention these metrics,

therefo  
ficult to  
pender  
racy. Ye  
or prec  
quantit  
unacce  
tive bel  
develop  
Stud  
include  
individ  
Several  
betwee  
any ind  
aged q  
points 1  
to 12

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



HELMHOLTZ  
| GEMEINSCHAFT

OPEN ACCESS

GFZ

Helmholtz-Zentrum  
POTSDAM

HELMHOLTZ  
| GEMEINSCHAFT



# Why Data Publication?

- Data driven research has become an important part of science.
- Scientific communication still emulates paper-based media.
- Most data remain inaccessible and are at risk of being lost.



# Data publication today

20 *B. Heim et al. / Global and Planetary Change 46 (2005) 9–27*

Table 6  
Overview on accuracies of chl-*a* algorithms (see also Table 4) applied on SeaWiFS data in July 2002 (07/20)

2002/07/20	HPLC	OC4	OC2	This study, July 2001+2002
<i>n</i> chl- <i>a</i> , all	22	17	17	17
<i>n</i> chl- <i>a</i> , case 1		17	17	17
Mean [ $\mu\text{g l}^{-1}$ ]	1.6	1.35	1.3	0.85
Median [ $\mu\text{g l}^{-1}$ ]	1.55	1.25	1.3	0.8
S.D. [ $\mu\text{g l}^{-1}$ ]	0.8	0.5	0.4	0.25
Accuracy, all [ $\mu\text{g l}^{-1}$ ]	$\pm 0.35$	$\pm 0.3$	$\pm 0.38$	$\pm 0.27\%$

2002/07/20	HPLC	Tuz et al. (2003), years 1994–1996	Tuz et al. (2003), year 1996 case 1	Gordon and Morel (1983)
<i>n</i> chl- <i>a</i> , all	22	17	17	17
<i>n</i> chl- <i>a</i> , case 1		17	17	17
Mean [ $\mu\text{g l}^{-1}$ ]	1.6	0.6	1	0.85
Median [ $\mu\text{g l}^{-1}$ ]	1.55	0.6	0.94	0.8
S.D. [ $\mu\text{g l}^{-1}$ ]	0.8	0.1	0.4	0.25
Accuracy, all [ $\mu\text{g l}^{-1}$ ]	$\pm 0.6$	$\pm 0.41$	$\pm 0.45$	$\pm 0.27\%$

Chl-*a* algorithms are OC2 (A, Table 4) and OC4 (B, Table 4), empirical chl-*a* algorithm (D, Table 4) from ground truth data set of Lake Baikal in 2001 and 2002 (this study), chl-*a* algorithms from Tuz et al. (2003); coefficient of studies from 1994 to 1996 (F, Table 4), coefficient of 1996 separately (G, Table 4), and case 1, Gordon and Morel (1983) (H, Table 4).

According to ground truth and SeaWiFS spectra for 2001–2002, the green peak of the highly transparent waters of Lake Baikal is commonly located SeaWiFS band 4 (510 nm). However, the absorbing and scattering optical activities in the presence of the terrigenous input shift the peak position towards SeaWiFS band 5 (555 nm). The waters in the observable cloud-free parts of the SeaWiFS acquisitions are not as turbid, so there does not occur a spectral shift in the peak position of the SeaWiFS spectra from SeaWiFS band 5 (555 nm) to band 6 (650 nm). This observed spectral behaviour of the peak shifting from 510 to 555 nm in the 2001–2002 SeaWiFS data sets of Lake Baikal can be simulated and reproduced using the bio-optical software ‘Water Colour Simulator’ (WASI) (Gege, 2004). This described spectral behaviour has been similarly shown from previous historical limnological studies. For example, Thomson and Jerome (1975) stated that clear waters of Lakes Ontario and Superior (USA) had a dominant wavelength of 490–530 nm, biologically more productive waters had a dominant wavelength of 550–560 nm, and waters with heavy sediment loadings had a dominant wavelength of >565 nm. This spectral shift is regarded as an indicator for the terrigenous input and can be used by applying a ‘mask of terrigenous input’ on the atmospherically corrected SeaWiFS data defined by reflectance ratio values of  $R_{RS510}/R_{RS555}$  below 0.9. This is in accordance to the SeaWiFS study done by Froidefond et al. (2002) in the Bay of Biscay, who observed chlorophyll overestimation (due to terrigenous input) in cases of  $R_{RS490}/R_{RS555}$  below 1. When calculating standard suspended matter products (Jørgensen, 2000; Binding et al., 2003), the high organic fluvial input in Barguzinski Bay and local fluvial input into the South Basin shows inverse grading with lowest calculated SPM concentrations towards the river inlets. Field spectrometer measurements and ground truth data show that, for several bio-optical processes of Lake Baikal, the assumption

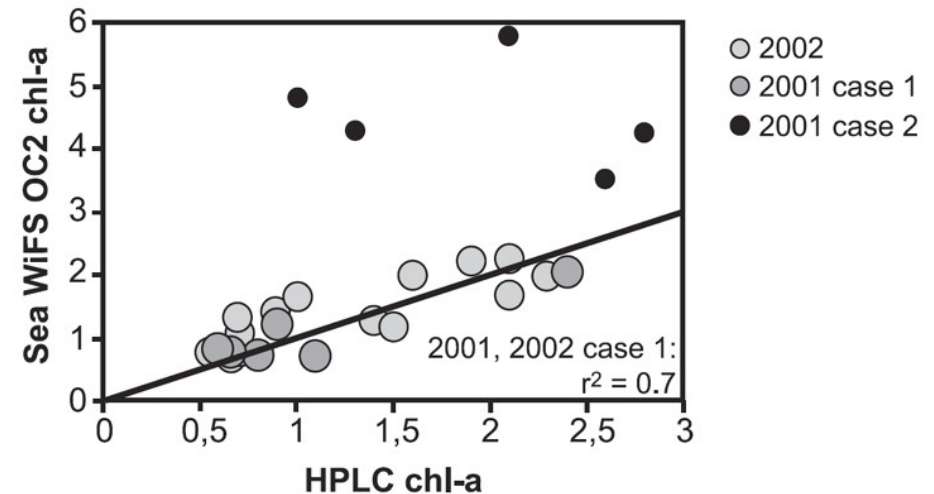
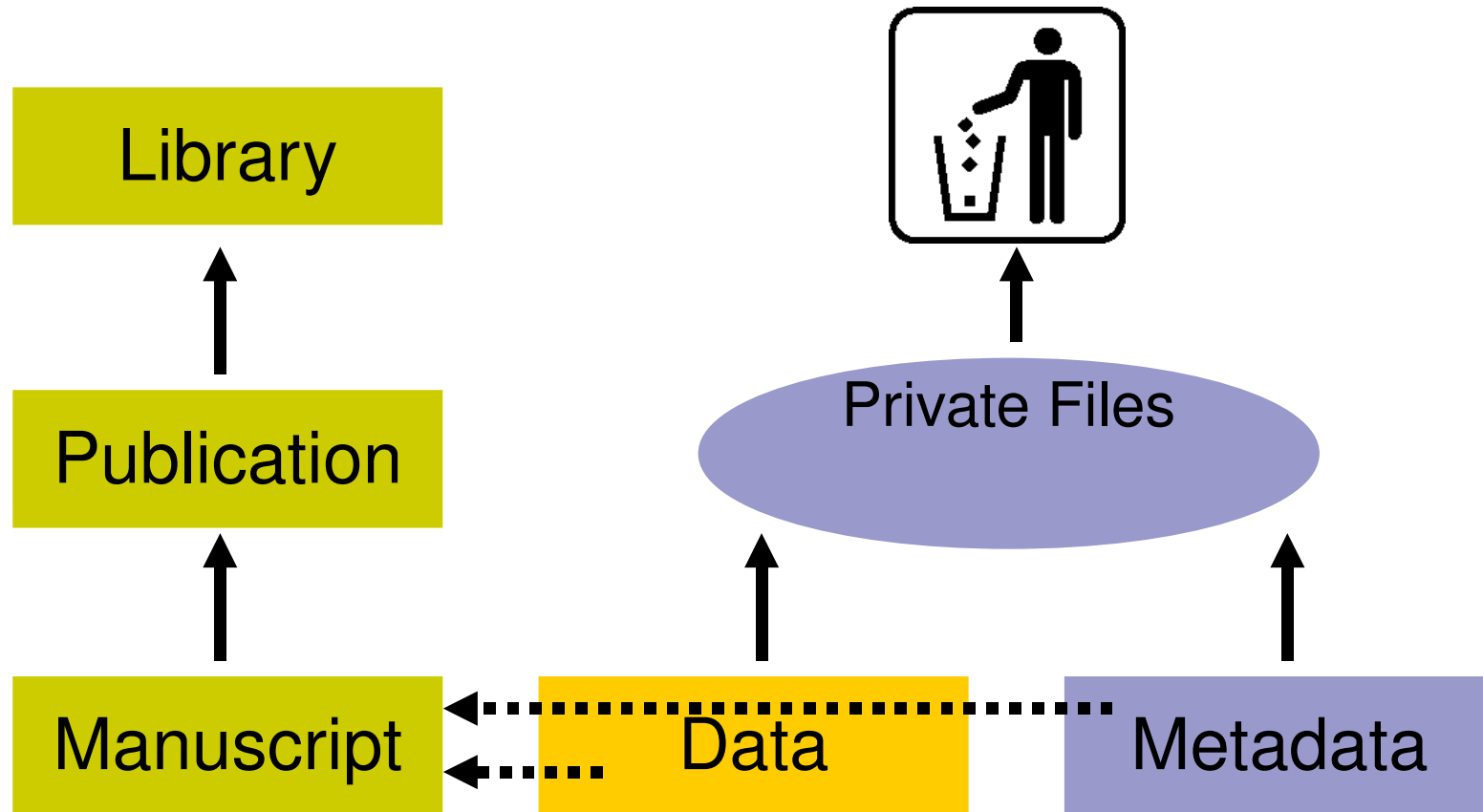
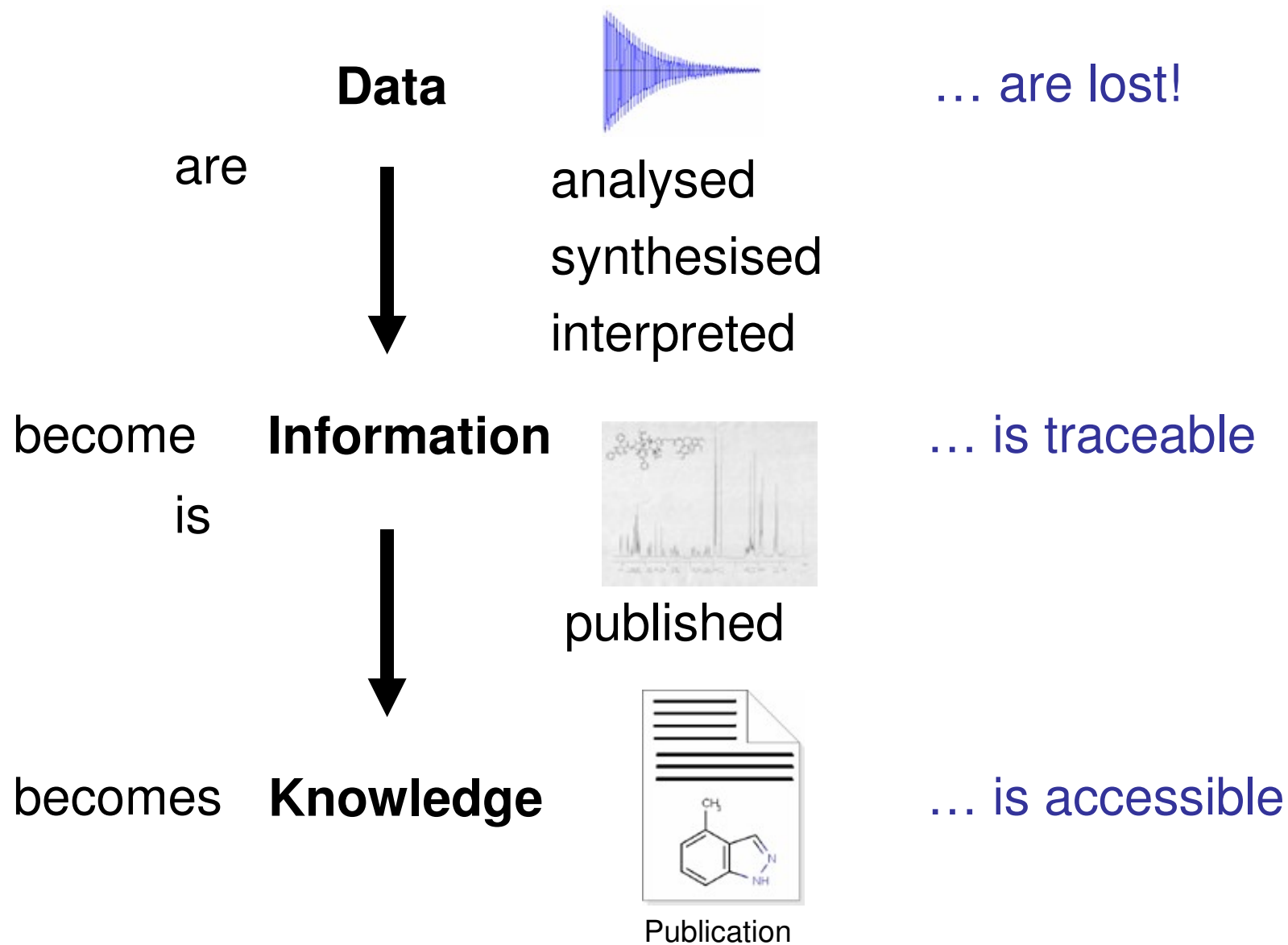


Fig. 2. The scattergram shows the relationship between concentrations of chl-*a* calculated from SeaWiFS OC2 and chl-*a* calculated determined from ground truth measurements during field expeditions in Lake Baikal during 2001 and 2002. Values of measured chlorophyll (HPLC) are the mean concentrations of each sampling point from 5 to 30 m depth. For the OC2 chl-*a* calculations, the most cloud-free acquisitions in 2001 (2001/07/19) and 2002 (2002/07/20) were chosen. Note the considerable chl-*a* overestimation caused by the influences of terrigenous input in case 2 waters. (Data available at: doi:10.1594/GFZ/EDP/CON/2004).

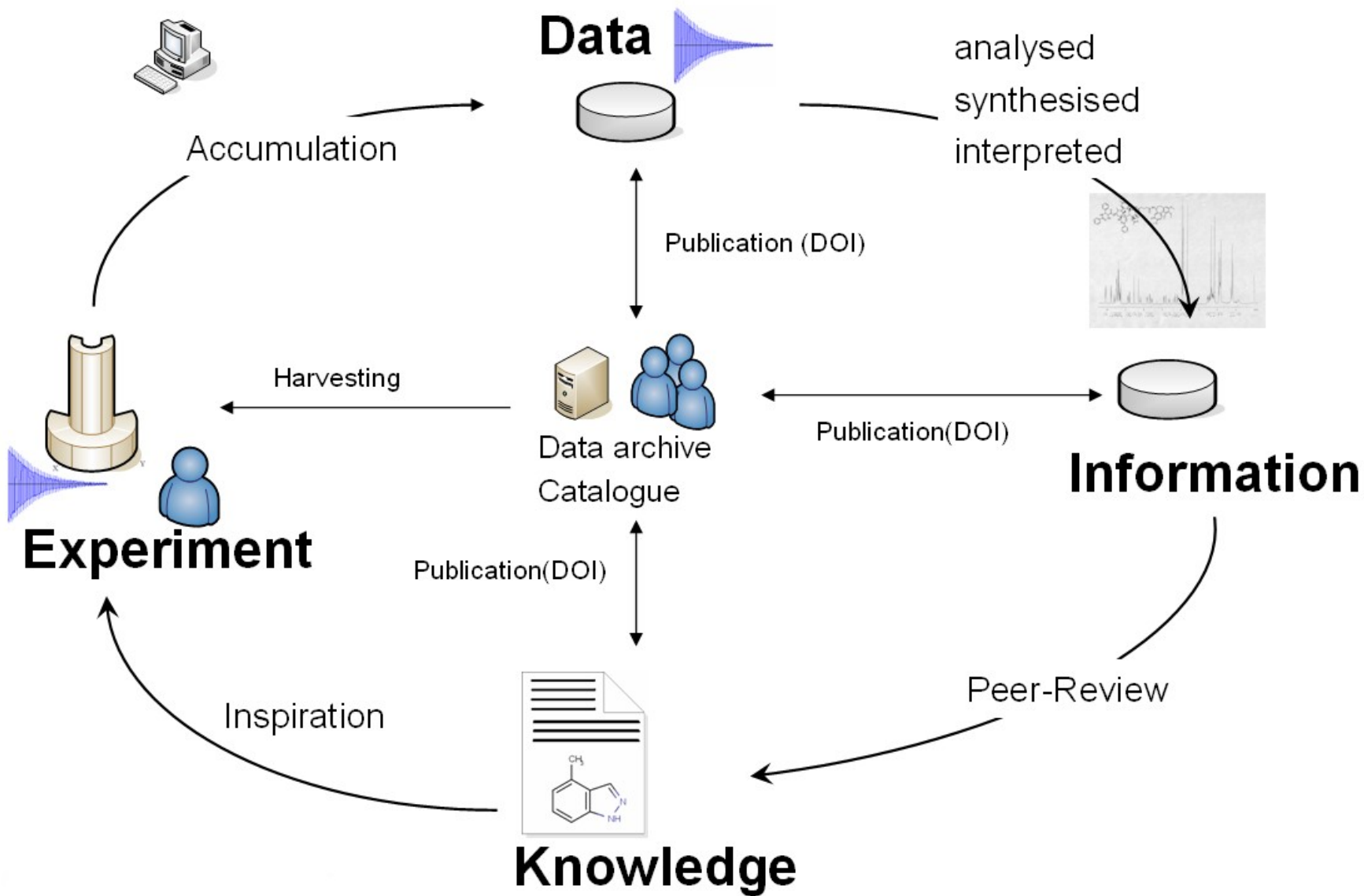
# Data in the publication process today



# The linear research trajectory



# The life-cycle of research



# Project “Publication and Citation of Scientific Primary Data”

- Funded by the German Science Foundation (DFG).
- Project partners:
  - TIB Hannover (Nat. Lib. Science and Technology)
  - WDC-MARE (Univ. Bremen/AWI Bremerhaven)
  - WDC Climate (MPI MET Hamburg)
  - GFZ Potsdam (proposed WDC-TERRA)
  - WDC-RSAT (DLR-DFD Oberpfaffenhofen)
- Implementation of services for the publication of data through “Publication Agents”.
- DOI registration at TIB Hannover, includes data publications into library catalogues.
- To date 18 DOI registration agents.

# DFG Project STD-DOI

- Projects „Publication and citation of scientific primary data“ (STD-DOI):
  - Make data citeable
  - Persistent Identifiers for Datasets (Digital Object Identifier, doi)
  - Methods for quality management in data publication.
- Data publication should (again) become part of common scientific practice.

# Making data citeable

- Putting data on the web is not new.
- How can data on the web be made citeable?
  - Unique name
  - Persistence
  - Globally resolvable location
- Prerequisites:
  - Globally unique persistent identifiers (Project STD-DOI)
  - Trusted archives (working group “Trusted Archives/Certification” in nestor and now DIN NABD15)



# Persistent Identifiers

- Everything is on the internet – you just can't find it.
- URLs have a half-life of only a few months. Not suitable for the record of science.
- A persistent identifier to a resource will provide ongoing access and reference to that resource and can be used whenever a permanent link is required.
- One of these systems is the Digital Object Identifier (DOI<sup>®</sup>).

# Name vs. Location

- A DOI gives a unique name to a digital object.
- The DOI can be resolved to a location (URL).
- The URL points to a node (web page) that describes the object and gives access to the object itself.
- Why a „splash page“?
  - More than one representation form
  - Access restrictions on grounds of policy
  - Access restrictions for technical reasons

# Persistence

- What makes a DOI persistent?
- Technically, DOI is a namespace of handle.net.
- Persistence is not only a matter of persistent technical systems, it is also a social construct.
- The members of the International DOI Foundation are committed to maintaining the persistence of the DOI system.
- The project „Publication and citation of primary scientific data“ chose DOI because it is an established and trusted brand.

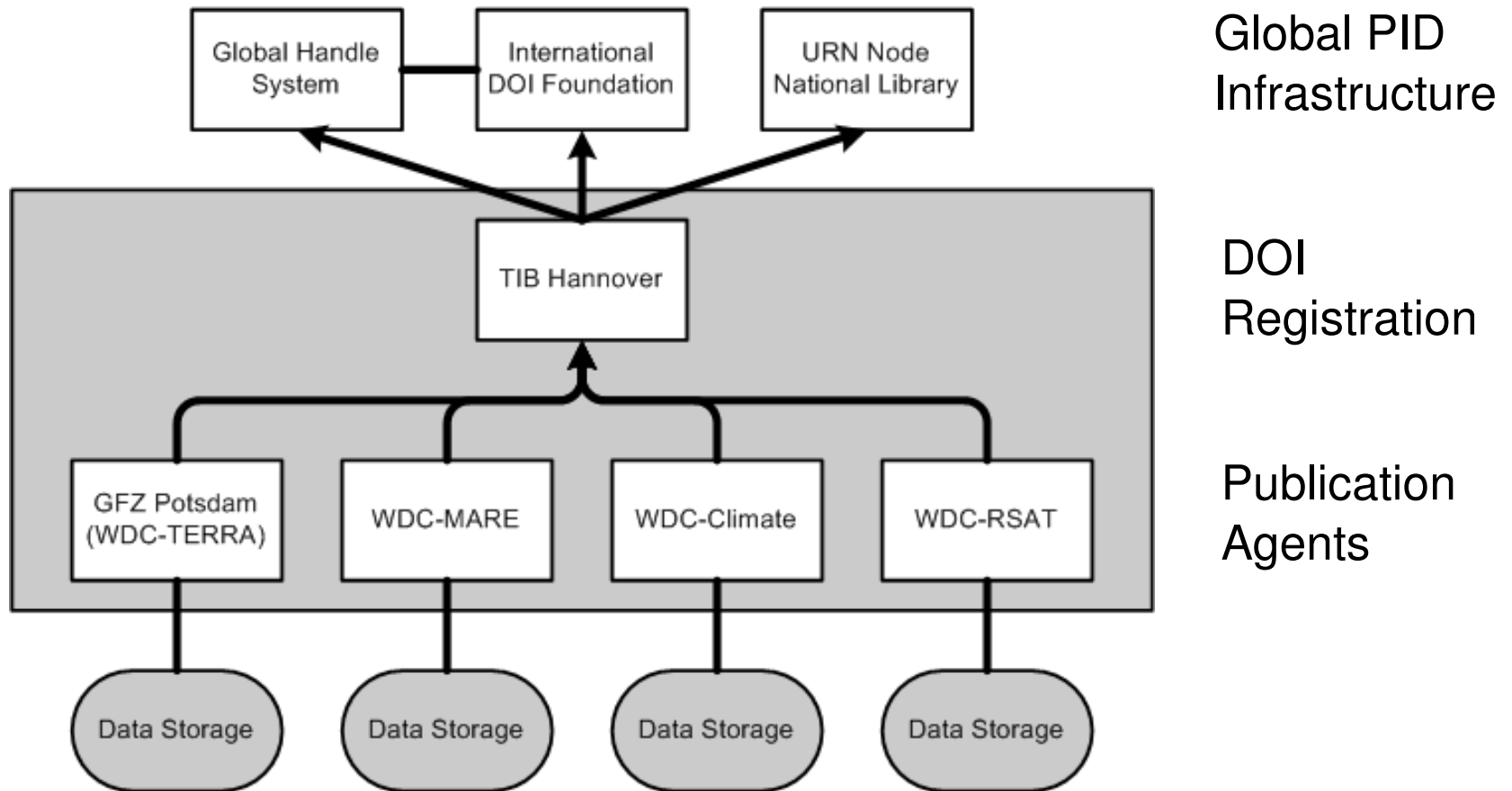
# Generation and Registration of DOIs

- DOIs are not issued by a central entity but only registered by TIB Hannover.
- Central concept: namespaces
- doi:10.1594/GFZ.SDDDB.1202
- doi:10. ... (IDF)
- doi:10.1594/ ... (TIB) = Publisher Prefix
- doi:10.1594/GFZ ... (GFZ)
- doi:10.1594/GFZ.SDDDB ... (SDDDB)
- doi:10.1594/GFZ.SDDDB.1202 (data item)

# Granularity

- What entities does a DOI identify?
- STD-DOI has two types of DOI:
  - Citation DOI (incl. GBV catalogue entry)
  - Data DOI (without GBV catalogue entry)
- DOI may also refer to collections of data objects.
- Example: doi:10.1594/PANGAEA.547983 refers to 426 datasets with DOIs (ozone sonde profiles from Georg-Forster-Station, Antarctica)
- Data granularity is case specific.

# STD-DOI System Architecture



# STD-DOI Metadata

- The STD-DOI metadata schema is fairly generic:
  - Bibliographical Data (author, publisher, Date, Abstract, ... )
  - Alternative web identifiers (URN, hdl, etc.)
  - Format and size
  - References to other Objects
- The reference to other objects may point to articles interpreting the data, other related data, physical samples, drill cores, ...

# Literature, Data, Samples



Search: ... **scholar** (0.09 seconds)

- ... [sediment distribution a](#) ... [Chilean continental slope related to upwelling and productivity](#) ... [Telegrafenberg - group of 2](#)
- D Hebbeln, M Marchant, T Freunberger, G Winter - Marine Geology, 2004 - ingentaconnect.com ... [coastal upwelling system of the Peru-Chile Current](#) ... [increase of surface water productivity](#) ... [known from ... about varying influences of upwelling](#) ... [upwelling](#)
- Evolution and biological effects of the 1997-98 El Niño in the upwelling ecosystem off northern ... [SFX Telegrafenberg - group of 3](#)
- O Uchiro, N Escobedo, H Homayouni, RA Quiroz ... - GEOPHYSICAL RESEARCH LETTERS, 2001 - profi.uoi.ac.il ... [The rich biological productivity of the Peru-Chile rise](#) ... [increases in productivity](#) ... [driven down coastal upwelling](#) ... [which brings colder, nutrient-rich, subsurface](#) ... [water](#) ... [upwelling](#)
- ... [Apaitze-otacoste associations off Peru and Chile](#) ... [paleo-oceanographic implications](#) ... [group of 2](#)
- VC Burnett, J Glead, Soc London, 1983 - ingentaconnect.com ... [to areas of intense coastal upwelling](#) ... [with associated high organic productivity](#) ... [in regions ... oceanographic/climatic trends in the Peru-Chile region](#) ... [upwelling](#)
- ... [Sediments below the Peru-Chile Current: controlling mechanisms and relationship with productivity](#) ... [group of 4](#)
- O Romero, D Hebbeln - Marine Microbiology, 2002 - ingentaconnect.com ... [surface sediments below the Peru-Chile Current: controlling mechanisms and relationship with productivity](#) ... [water production](#) ... [coastal upwelling](#) ... [off northern ...](#)
- High- and low-latitude climate control on the position of the southern Peru-Chile Current during the ... [group of 5](#)
- F Lamy, C Rühlemann, D Hebbeln, G Weller - Paleoceanography, 2002 - ingentaconnect.com ... [of the world, the strong advection and upwelling of cold ...](#) ... [Among the El Niño](#) ... [Peru-Chile Cur-rent \(PCC ... km\)](#) ... [and an exceptionally high productivity](#) ... [supporting an ...](#)
- ... [Peru Upwelling Region Sediments Near 15°S](#) ... [Dissolved Free and Total Hydrolyzable Amino Acids](#)
- SM Henrichs, JW Farrington, C Lee - Limnology and Oceanography, 1984 - JSTOR ... [The Peru upwelling region is ... part of regional upwelling ...](#) ... [productivity and the ...](#) ... [in sulphide biota under Peru-Chile subsurface](#) ... [concentration](#) ... [upwelling](#)
- ... [Seasonal variations of the particle flux in the Peru-Chile current at 30 S under "normal" ...](#)
- D Hebbeln, M Marchant, G Winter - Deep Sea Res., Part 1, 2000
- Peru Upwelling Region Sediments Near 15°(S). 1. Remineralization and Accumulation of Organic Matter - group of 2
- SM Henrichs, JW Farrington - Limnology and Oceanography, 1984 - JSTOR

doi:10...

Table 4. Overview on average ...

Year	SPIC	OC2	OC1	OC4	OC6
2001	11	17	17	17	17
2002	11	17	17	17	17
2003	14	13	13	13	13
2004	14	13	13	13	13
2005	6	6	6	6	6
2006	7	17%	17%	17%	17%

Earth System Science

# Data

The International Geosphere-Biosphere Programme

Volume 1 | Number 1 | 2008

Copenhagen Publications

Sref: ...

doi:10.1594/...

Scientific Drilling Database

Home | About SIO | News | Data Publications | Catalogue | Search | Admin

Activities: COMB1501

Latitude: 52.9667  
Longitude: 107.8900  
Elevation: -1260  
DateTime: 2001-07-18 03:53:40  
Program: High-resolution CONTINENTAL paleoseismic record in Lake Baikal  
ExpDepth: C0401-6  
Platform: BA-Verschiebung  
Gear: Water sample

IGSN hdl: ...

doi:10...

Table 4. Overview on average ...

Year	SPIC	OC2	OC1	OC4	OC6
2001	11	17	17	17	17
2002	11	17	17	17	17
2003	14	13	13	13	13
2004	14	13	13	13	13
2005	6	6	6	6	6
2006	7	17%	17%	17%	17%





# Peer-review of data

- Peer-review of data should be common practice in scholarly communication.
- Present day practice varies widely between disciplines.
- Aspects of data peer-review:
  - Material
  - Methods
  - Documentation
  - Plausibility
- Guiding principle: can the data be verified and re-used?

# Syntactical Checks

- Some disciplines have already developed syntactical checks of data.
- Examples:
  - Expected range of values
  - Expected range of errors
  - Distribution of first digits (Poisson distribution)
  - Distribution of last digits (even distribution)

# Use cases

The project STD-DOI defines three use cases:

- Data publication as supplement to a journal publication
- Data publication as an independent work
- Data publication from data streams and by automated systems (e.g. environmental monitoring, earth observation, observatories).

# Scientific Drilling Database

Data from Deep Earth Sampling and Monitoring

- + Home
- + About SDDB
- + News
- + Data Publications

## Dataset Description

Citation: [Heim, Birgit; Oberhänsli, Hedi; Fietz, Susanne; Kaufmann, Hermann; \(2006\): The relationship between concentrations of chl-a calculated from SeaWiFS OC2 and chl-a calculated determined from ground truth measurements during field expeditions in Lake Baikal during 2001 and 2002](#)



Glossary  
Catalogue

Citation: [Heim, Birgit; Oberhänsli, Hedi; Fietz, Susanne; Kaufmann, Hermann; \(2006\): The relationship between concentrations of chl-a calculated from SeaWiFS OC2 and chl-a calculated determined from ground truth measurements during field expeditions in Lake Baikal during 2001 and 2002. \*Scientific Drilling Database\*. doi:10.1594/GFZ.SDDB.1043](#)

[Download Citation \(EndNote\)](#)

during 2001 and 2002

Related Publications:

- [Birgit Heim, Hedi Oberhaensli, Susanne Fietz and Hermann Kaufmann, Variation in Lake Baikal's phytoplankton distribution and fluvial input assessed by SeaWiFS satellite data, Global and Planetary Change, Volume 46, Issues 1-4, Progress towards reconstruct doi:10.1016/j.gloplacha.2004.11.011](#)

[Show in Google Earth](#)

Related Publications:

- [Birgit Heim, Hedi Oberhaensli, Susanne Fietz and Hermann Kaufmann, Variation in Lake Baikal's phytoplankton distribution and fluvial input assessed by SeaWiFS satellite data, Global and Planetary Change, Volume 46, Issues 1-4, Progress towards reconstruct doi:10.1016/j.gloplacha.2004.11.011](#)

Activities:

**CON01-501-1**

Latitude: 52.6667 °N



# Procedure

- Peer-review as part of the editorial process of the journal, including review of the underlying data.
- Formal checks (syntactic)

# Data publication as independent work

You are not logged in (LOG IN)



Earth System Science Data  
The Data Publishing Journal

| Copernicus.org |

[Volumes and Issues](#) [Contents of Issue 1](#)

[Home](#)

[Online Library ESSD](#)

[Online Library ESSDD](#)

- [Papers in Open Discussion](#)
- [Volumes and Issues](#)
- [Special Issues](#)
- [Library Search](#)
- [Title and Author Search](#)

[General Information](#)

[Submission](#)

[Review](#)

[Production](#)

[Subscription](#)

[Comment on a Paper](#)

ARCHIVED IN



PORTICO

Earth Syst. Sci. Data Discuss., 1, 1-13, 2008

[www.earth-syst-sci-data-discuss.net/1/1/2008/](http://www.earth-syst-sci-data-discuss.net/1/1/2008/)

© Author(s) 2008. 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. Germandt**

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

**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.

[Discussion Paper](#) (PDF, 423 KB) [Interactive Discussion](#) (Final Response, 7 Comments)

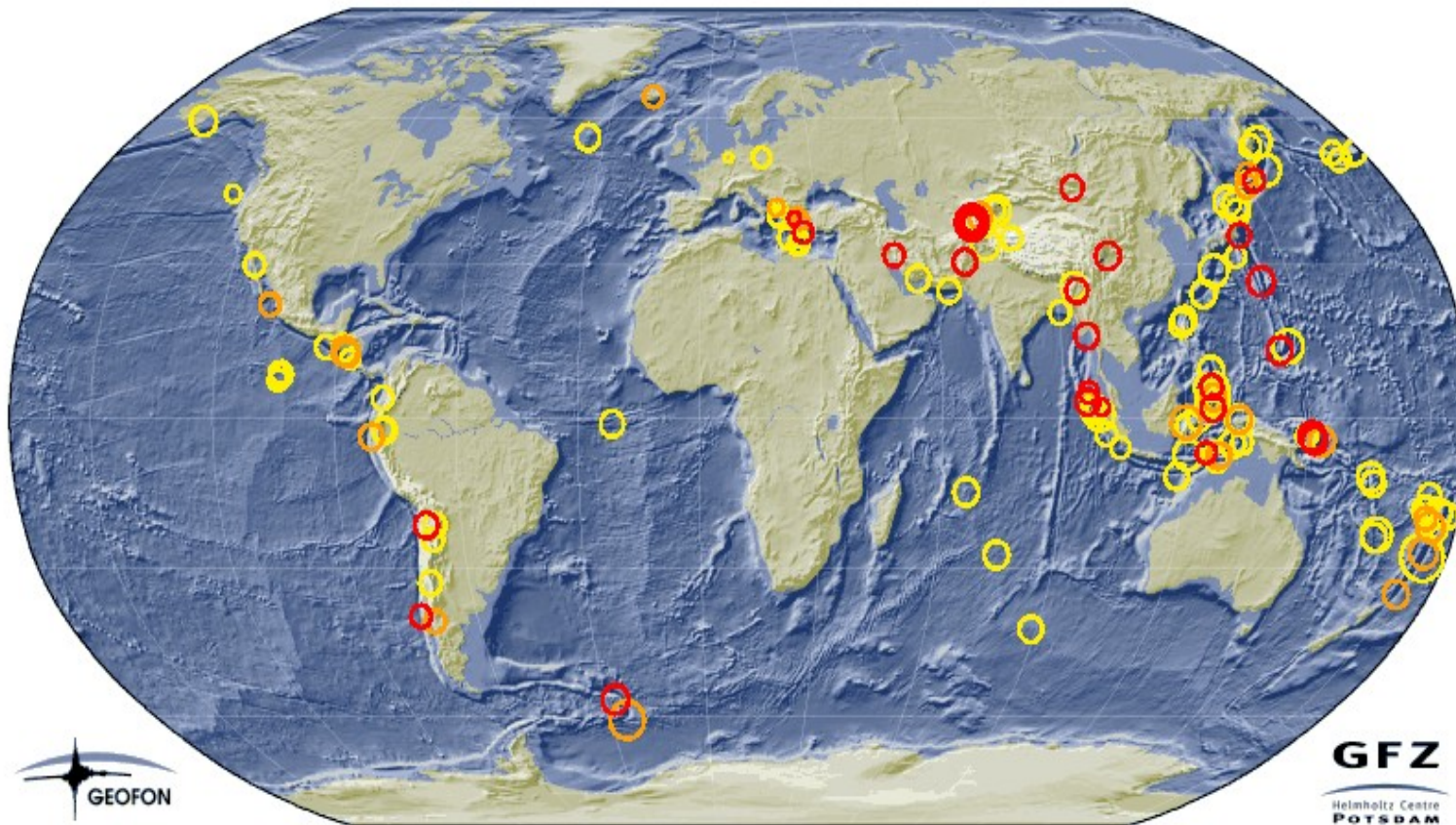
**Citation:** König-Langlo, G. and Germandt, H.: Compilation of ozonesonde profiles from the Antarctic Georg-Forster-Station from 1985 to 1992, Earth Syst. Sci. Data Discuss., 1, 1-13, 2008. [Bibtex](#) [EndNote](#) [Reference Manager](#)

# Procedure

- Peer-review by data journal (e.g. Earth System Science Data, G-Cubed Data Briefs)
- Formal checks (syntactic)

# Data from automated systems

## Automatic GEOFON Global Seismic Monitor



The displayed events are within the last **24 hours** / **1-4 days** / **4-14 days**.



Most recent large event:

**Afghanistan Tajikistan Border Region**

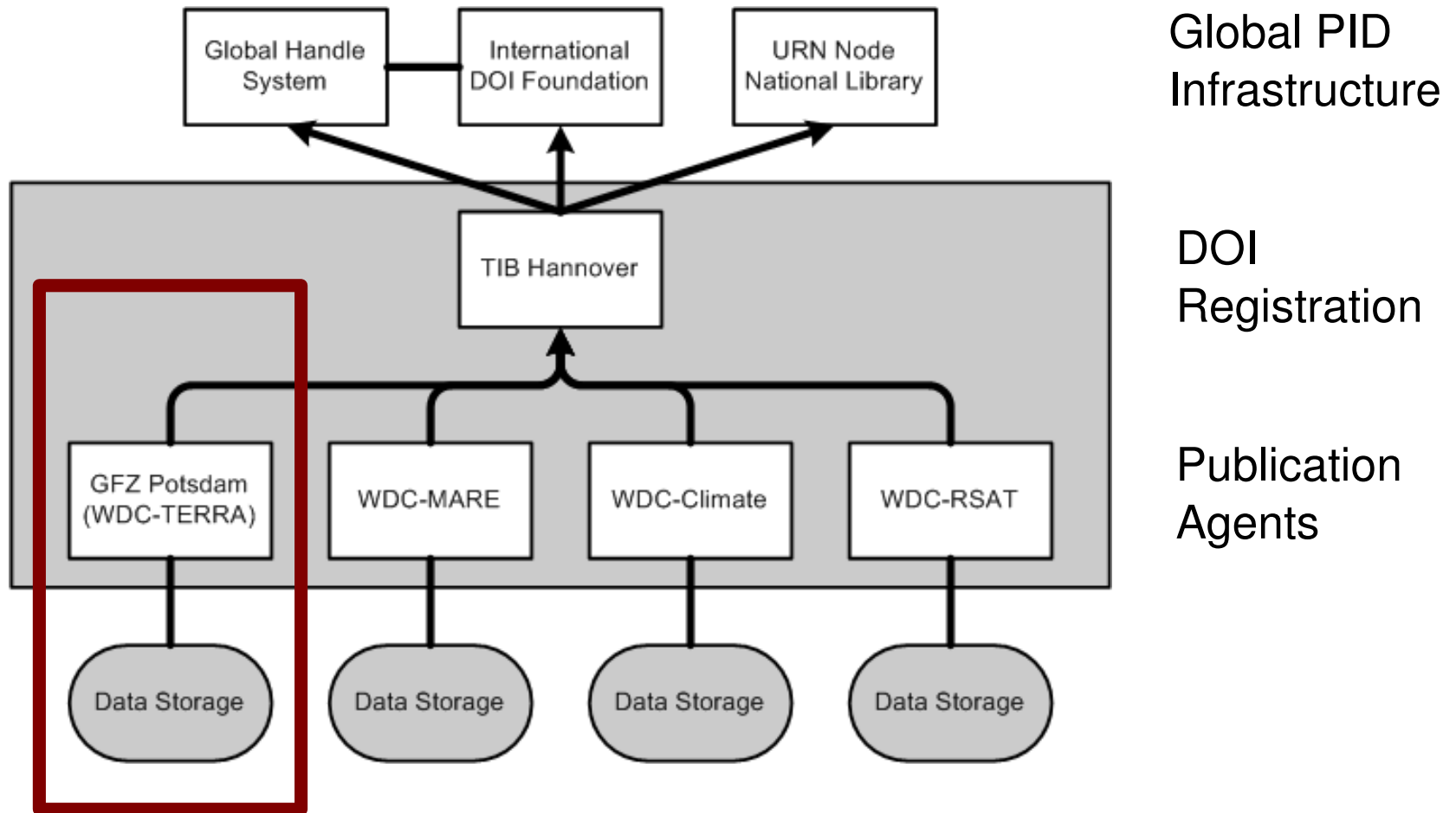
**Magnitude: 5.0**



# Procedure

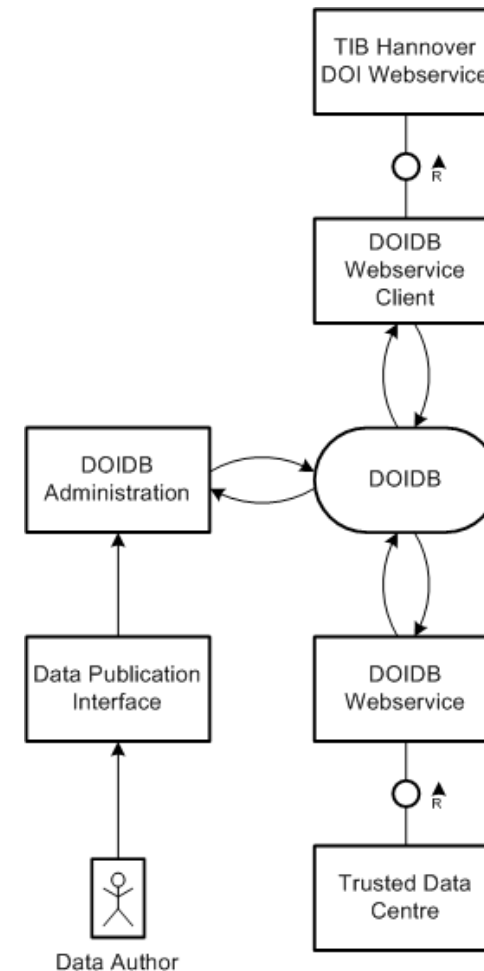
- Peer-review of the observatories and processing chains
- Quality management of the processing chains

# STD-DOI System Architecture



# DOIDB

- GFZ publishes data from more than one source.
- The web service call to the TIB DOI service is moderated and logged by the GFZ DOIDB service.
- Use cases:
  - Automatic registration by Trusted Agent
  - Manual registration for single data sets.



# More on data publication in practice

- The STD-DOI project has modelled data publication to be as close as possible to the established modes of scientific publication.
- Governing principles:
  - Stable record of science
  - Unambiguousness
  - Adequacy
  - Quality assured data

# Example: Dataset has been changed

- Dataset has been changed due to update:
  - New dataset is published alongside with the existing dataset as new **edition** with its own DOI.
  - Datasets refer each other by their „related object“ elements.
- Dataset has been changed due to corrections:
  - New dataset is published alongside with the existing dataset as **erratum** with its own DOI.
  - Datasets refer each other by their „related object“ elements.

# Example: Dataset does no longer exist

- If a dataset no longer exists its DOI resolves to a web page that explains the reasons why the published object is no longer available.

# Example: growing datasets

- In the case of time series from environmental monitoring, the data set continually grows through data being added later in time.
  - Time series is identified by a DOI.
  - Subsets can be identified by DOI or other identifiers
  - Subsets can be identified by giving a canonical path (here: time and space coordinates).

# Canonical Path

- A canonical path is a standardised description of the path leading to a specific sub-element.
- Example: Genesis 1:1
  - Collection: The Bible
  - Book: Genesis
  - Chapter: 1
  - Verse: 1

(Genesis 1:1) In the beginning God created the heaven and the earth.



# Canonical path

Variations of the earth magnetic field over South Africa in 2008:

- Rother, Martin (2000): CH-ME-2-FGM-NEC corrected magnetic field in ECEF/NEC system
- doi:10.1594/GFZ.ISDC.CHAMP/CH-ME-2-FGM-NEC
- Start date: 2008-01-01 End date: 2008-12-31
- Bounding box: NW corner 23° S 16° E, SW corner 35° S 33° E

# Data and Copyright

- Copyright law sees data as unoriginal. Collections of facts cannot be copyrighted.
- „Sweat of the brow“ does not protect a data collection.
- The EU directive on databases protects databases to a certain extent.
- Workaround: End-user licence agreements.
- The legal framework for research data leaves much to be desired.
- GFZ recommends CC-by licence.

# The Future

- STD-DOI has become a service of TIB Hannover and will become an international service on 2009-12-01 (DataCite).
- More and more publication agents are joining the system.
- Talks have been held with journal editors to make data publication part of the journal policy.
- Publishing houses to add data to their value added services.

# Linking literature to data

Home Browse Search My Settings Alerts Help

Quick Search All fields  Author

search tips Journal/book title  Volume  Issue  Page  Clear Go Advanced Search

Marine Micropaleontology  
Volume 66, Issues 3-4, 20 February 2008, Pages 208-221

Font Size:

Abstract

doi:10.1016/j.marmicro.2007.10.002  
 Cite or Link Using DOI

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

## Centennial-scale climate variability in the Timor Sea during Marine Isotope Stage 3

Anke Dürkop<sup>a</sup>, , Ann Holbourn<sup>a</sup>, Wolfgang Kuhnt<sup>a</sup>, Rina Zuraída<sup>a, b</sup>, Nils Andersen<sup>c</sup> and Pieter M. Grootes<sup>c</sup>

<sup>a</sup>Institute of Geosciences, Christian-Albrechts-University, Ludwig-Meyn-Str. 10-14, D-24118 Kiel, Germany

<sup>b</sup>Leibniz-Institute of Marine Sciences, IFM-GEOMAR, Wischhofstr. 1-3, D-24148 Kiel, Germany

<sup>c</sup>Leibniz-Laboratory for Radiometric Dating and Stable Isotope Research, Christian-Albrechts-University, Max-Eyth-Str. 11 - 13, D-24118 Kiel, Germany

Received 4 June 2007; revised 1 October 2007; accepted 4 October 2007. Available online 18 October 2007.

### Abstract

We present a high-resolution (~ 60–110 yr) multi-proxy record spanning Marine Isotope Stage 3 from IMAGES Core MD01-2378 (13°04.95' S and 121°47.27' E, 1783 m water depth), located in the Timor Sea, off NW Australia. Today, this area is influenced by the Intertropical Convergence Zone, which drives monsoonal winds during austral summer and by the main outflow of the Indonesian Throughflow, which represents a key component of the global thermohaline circulation system. Thus, this core is ideally situated to monitor the linkages between tropical and high latitude climate variability. Benthic  $\delta^{18}\text{O}$  data (*Planulina wuellerstorfi*) clearly reflect Antarctic warm events (A1–A4) as recorded by the EPICA Byrd and Dronning Maud Land ice cores. This southern high latitude signal is transferred by deep and intermediate water masses flowing northward from the Southern Ocean into the Indian Ocean. Planktonic  $\delta^{18}\text{O}$  shows closer affinity to northern high latitudes planktonic and ice core records, although only the longer-lasting Dansgaard–Oeschger warm events, 8, 12, 14, and 16–17 are clearly expressed in our record. This northern high latitude signal in the surface water is probably transmitted through atmospheric teleconnections and coupling of the Asian–Australian monsoon systems. Benthic foraminiferal census counts suggest a coupling of Antarctic cooling with carbon flux patterns in the Timor Sea. We relate increasing abundances of carbon-flux sensitive species at 38–45 ka to the northeastward migration of the West Australian Current frontal area. This water mass reorganization is also supported by concurrent decreases in Mg/Ca and planktonic  $\delta^{18}\text{O}$  values (*Globigerinoides ruber* white).

**Keywords:** Marine Isotope Stage 3; stable isotopes; foraminifera; Timor Sea; paleoclimate

### Article Outline

1. Introduction
2. Material and method

### Article Toolbox

- E-mail Article
- Cited By
- Save as Citation Alert
- Citation Feed
- Export Citation
- Add to my Quick Links
- Permissions & Reprints
- Cited By in Scopus (2)

### Related Articles in ScienceDirect

- Evidence for Millennial-Scale Climate Change During Marine Isotope Stage 3  
*Quaternary Research*
  - Climate variability and land-ocean interactions in the last glacial period  
*Marine Micropaleontology*
  - QUATERNARY STRATIGRAPHY | Climatostratigraphy  
*Encyclopedia of Quaternary Science*
  - Changes in the thermocline structure of the Indonesian Ocean during Marine Isotope Stage 3  
*Earth and Planetary Science Letters*
  - What is what in the ice and the ocean?  
*Quaternary Science Reviews*
- [View More Related Articles](#)

**Supplementary Data**  
[SFX Telegrafenberg](#)  
[View Record in Scopus](#)

# DataCite

- A group of leading research libraries and technical information providers have established a partnership to improve access to research data on the internet.
- The goal of DataCite is to establish a not-for-profit agency that enables organisations to register research datasets and assign persistent identifiers.
- Founding ceremony of DataCite will be held at the German embassy in London on 2009-12-01.

# DataCite

- **Members:**
  - TIB Hannover (German Nat. Lib. Sci. Tech.)
  - British Library
  - ETH Zürich
  - French Institute for Scientific and Technical Information (INIST)
  - Technical Information Center of Denmark
  - Library Technical University Delft
  - Canada Institute for Scientific and Technical Information (CISTI)
  - Australian National Data Service (ANDS)
- DataCite is open to new members.

# Thank you!

For more information go to

<http://www.std-doi.org>

<http://www.datacite.org>