

A system for digital preservation of scientific data using DOI Names

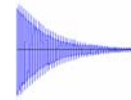
Uwe Rosemann

International Conference on Preservation of Digital Objects
Beijing, China

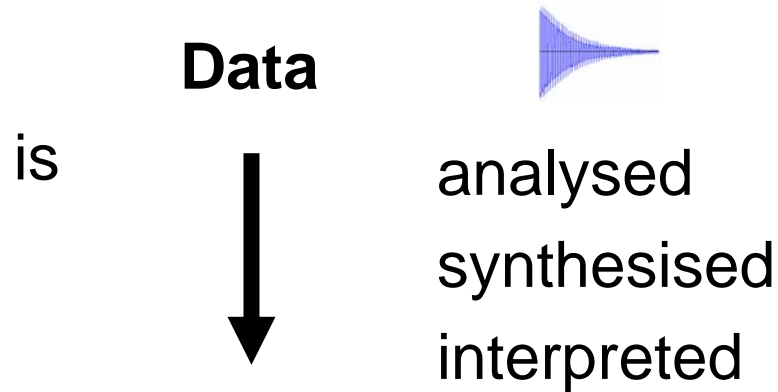


Problem: The research trajectory

Data

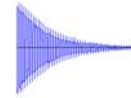


Problem: The research trajectory



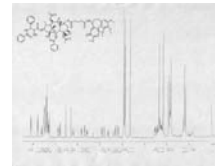
Problem: The research trajectory

Data

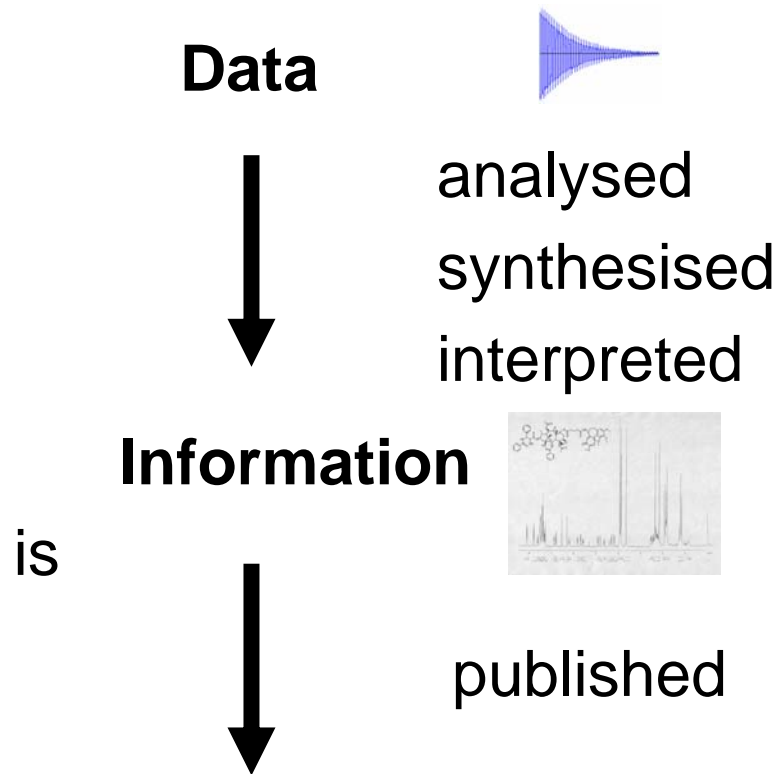


analysed
synthesised
interpreted

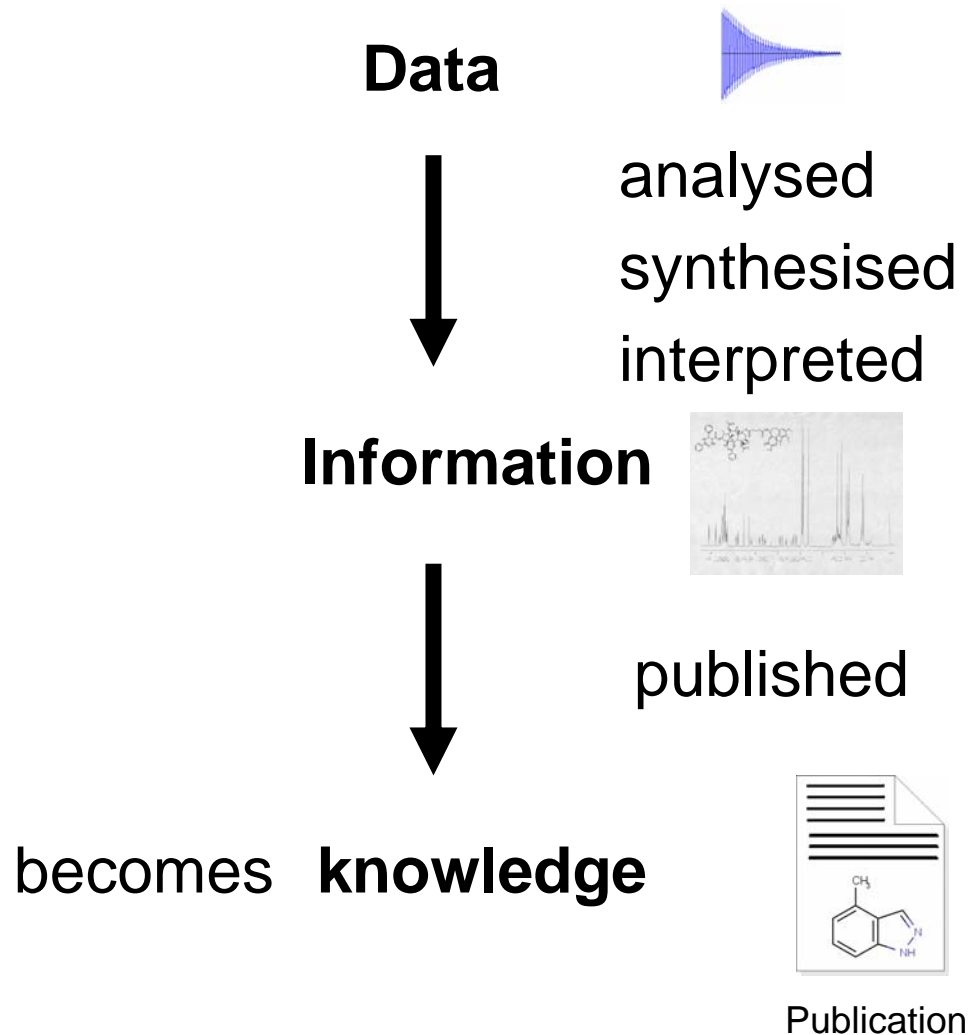
becomes **Information**



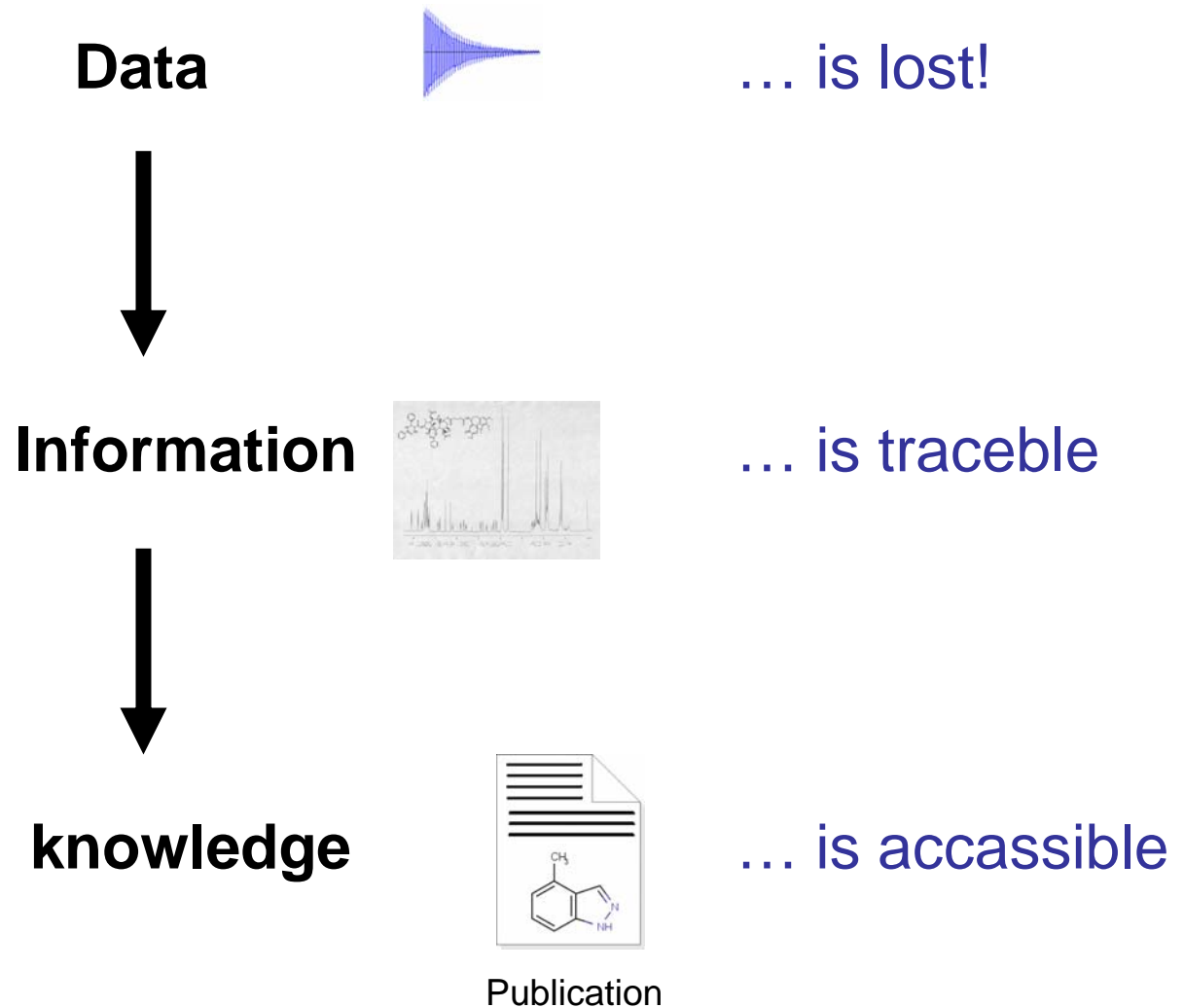
Problem: The research trajectory



Problem: The research trajectory



Problem: The research trajectory



This leads to...

- > ... unnecessary duplication of research efforts
- > The verification of results becomes difficult
- > Large amounts of research funds are spent every year, while already existing data remain underutilised

We need..

- ... a persistent identifier that allows global access to the data sets and all relevant metadata,
- To enabling citations of data
- To encourage good scientific practice
- To acknowledge scientific work

Project

- The *German Research Foundation* (DFG) has started the project *Publication and Citation of Scientific Primary Data* to increase the accessibility of scientific primary data, starting with the field of earth science.
- The *German National Library of Science and Technology* (TIB) is now established as a “non-commercial” DOI-registration agency for scientific primary data as a member of the *International DOI Foundation* (IDF).

Primary data publication

- During her research for the *World Data Center Climate (WDCC)* the scientist Mrs. Weather gains primary data about the temperature in Hannover in the year 2003.
- As usual the primary data is tested, evaluated, stored and administrated at the *WDCC*.
- **In addition** Mrs. Weather registers the primary data at the **TIB**.

Registration of primary data

Mrs. Weather transmits to the TIB the URL where the data can be accessed, together with a XML-file containing all relevant bibliographic metadata

Including all information obligatory for the citing of electronic media (ISO 690-2)

- author
- title
- size
- edition
- language
- publisher
- publishing date
- publishing place

Identifier

> **The TIB includes the metadata about the primary data in her catalogue, and awards the primary data with a unique identifier for registration: a DOI name**

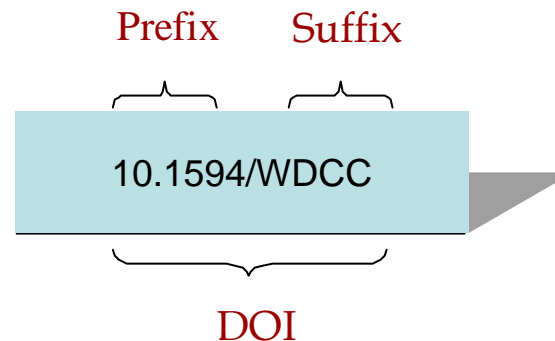
In her publications, Mrs. Weather is now citing this primary data with its unique DOI name:

Weather, Hermione (2003): Temperature in Hannover for 2003.

[doi: 10.1594/WDCC/W_Han_2003_MMB_2]

Digital Object Identifier (DOI)

- The DOI system is a worldwide system for persistent and actionable identification and interoperable exchange of intellectual property on digital networks;
- A DOI name is made up of two components, the prefix and the suffix;



- The value of the DOI system lies in its combination of Resolution, Metadata and Policy.

Resolving the DOI

This DOI name can be resolved (and the data can be accessed) in every browser worldwide in different ways:

- http://dx.doi.org/10.1594/WDCC/W_Han_2003_MMB_2
- http://doi.tib-hannover.de:8000/10.1594/WDCC/W_Han_2003_MMB_2

Or by

[Doi://10.1594/WDCC/W_Han_2003_MMB_2](https://doi.org/10.1594/WDCC/W_Han_2003_MMB_2)

(after installing a browser plugin)

Scope

The TIB registers primary data worldwide from a scientific, technical or medical background.

So far over 450,000 datasets have received a DOI name.

Furthermore TIB registers any scientific or technical content that is a result of community funded research in Europe.

The storage, maintaining and evaluation of the contents lies with the content providers (institutions, data centers, ...) as so-called publication agents.

Publication agent

- > **Content providers, Data centers**
- > **Who are responsible for:**
 - > Quality assurance
 - > Storage of the content and accessibility
 - > Creation of metadata
- > **The DOI registration can be included into the Workflow of the publication agent**
- > **The TIB stores the metadata and keeps it searchable.**

sort by
User id: | [logout](#)
[Title list](#) | **[title data](#)**

ve search results

Databases

Order without

search

Profile

TIB Homepage

 ■ **results** search [and] (all words) Yancheva

1 of 2



Title: [Rock magnetism](#) and [X-ray fluorescence spectrometry analyses](#) on [sediment cores](#) of the [Lake Huguang Maar](#), [Southeast China](#) [[supplementary data](#) to the [reference given](#)]

Author: [Gergana Yancheva](#) ; [Norbert R Nowaczyk](#) ; [J Mingram](#) ; [Peter Dulski](#) ; [Georg Schettler](#) ; [Jörg F W Negendank](#) ; [Jiaqi Liu](#) ; [Daniel M Sigman](#) ; [Larry S Peterson](#) ; [Gerald Haug](#)

Published: Bremen/Bremerhaven : PANGAEA - Publishing Network for Geoscientific & Environmental Data, 2007-01-23

Extent: 4 Datasets.

Note: This dataset is cited by doi: 10.1038/nature05431

Abstract: The Asian-Australian monsoon is an important component of the Earth's climate system that influences the societal and economic activity of roughly half the world's population. The past strength of the rain-bearing East Asian summer monsoon can be reconstructed with archives such as cave deposits, but the winter monsoon has no such signature in the hydrological cycle and has thus proved difficult to reconstruct. Here we present high-resolution records of the magnetic properties and the titanium content of the sediments of Lake Huguang Maar in coastal southeast China over the past 16,000 years, which we use as proxies for the strength of the winter monsoon winds. We find evidence for stronger winter monsoon winds before the Bølling-Allerød warming, during the Younger Dryas episode and during the middle and late Holocene, when cave stalagmites suggest weaker summer monsoons. We conclude that this anticorrelation is best explained by migrations in the intertropical convergence zone. Similar migrations of the intertropical convergence zone have been observed in Central America for the period AD 700 to 900, suggesting global climatic changes at that time. From the coincidence in timing, we suggest that these migrations in the tropical rain belt could have contributed to the declines of both the Tang dynasty in China and the Classic Maya in Central America.

REFERENCE:

Yancheva, Gergana; Nowaczyk, Norbert R; Mingram, J; Dulski, Peter; Schettler, Georg; Negendank, Jörg F W; Liu, Jiaqi; Sigman, Daniel M; Peterson, Larry S; Haug, Gerald (2007): Influence of the intertropical convergence zone on the East Asian monsoon, Nature, 445, 74-77, doi:10.1038/nature05431

Techn. data: Format: application/zip

Links: doi: [10.1594/PANGAEA.587840](https://doi.org/10.1594/PANGAEA.587840)

URN: [urn:nbn:de:tib-10.1594/PANGAEA.5878400](https://nbn-resolving.org/urn:nbn:de:tib-10.1594/PANGAEA.5878400)

Holding: [Display free access!](#)

Note: Primaerdaten

1 of 2

Always quote citation when using data!

DOI for Scientific and Technical Data

10.1594/WDCC/EH5-T63L31_OM-GR1.5L40_CTL_6H

Title

IPCC-AR4 MPI-ECHAM5_T63L31 MPI-OM_GR1.5L40 PIntrl(pre-industrial control experiment): atmosphere 6 HOUR values MPImet/MaD Germany

Citation

Roeckner, Erich; Lautenschlager, Michael; Esch, Monika 2006; IPCC-AR4 MPI-ECHAM5_T63L31 MPI-OM_GR1.5L40 PIntrl(pre-industrial control experiment): atmosphere 6 HOUR values MPImet/MaD Germany. [doi: 10.1594/WDCC/EH5-T63L31_OM-GR1.5L40_CTL_6H]

Publication Date

2006-06-29

Author(s)

Roeckner, Erich; Lautenschlager, Michael; Esch, Monika

Summary

The data represent 6 hourly values of pre-industrial climate simulation. Here, for the year 1860, concentrations of well mixed greenhouse gases (CO₂, CH₄, N₂O) were prescribed. Ocean data (MPI-OM) are available within this experiment, they are stored in EXTRA FORMAT. Datasets with monthly mean values are also available.

As the boundary conditions are not time dependent, the time access is arbitrary.

Technical data to this experiment:

The experiment is using ECHAM5.2.02 coupled to MPI-OM Vers.1.0 GR1.5L40 and was run on a NEC-SX6(hurrikan).

The output from the model run: hurrikan.dkrz.de:/ut/3/m214002/EXP500/run520

Location(s)

World; Latitude: -90.0 to 90.0; Longitude: 0.0 to 360.0

Spatial Coverage

Latitude: -90.0 to 90.0; Longitude: 0.0 to 360.0; Altitude: -6.98 m to 10.0 hPa

Temporal Coverage

1/1/2150 - 31/12/2655 (calendrical, arbitrary numbered years)

Data Format(s)

GRIB

Datasize

8310350578836 Bytes

Contact

Dr. Erich Roeckner

Max-Planck-Institut fuer Meteorologie

Bundesstrasse 53, D-20146 Hamburg, F.R. Germany

<http://www.mpimet.mpg.de/>

Project

IPCC-Hamburg Climate Model Simulation (IPCC_HH)

The Intergovernmental Panel on Climate Change (IPCC) has been established by WMO and UNEP to assess scientific, technical and socio-economic information, relevant for the understanding of climate change, its potential impacts and options for adaption and mitigation.

Continued description about the work of the IPCC will be found at the homepage (<http://www.ipcc.ch>) and (www.grida.no/climate/ipcc). As a further development the Special Report on Emission Scenarios (SRES) have been constructed, to explore future developments in the global environment with special reference to the production of greenhouse gases and aerosol precursor emissions.

A set of four scenarios families (A1, A2, B1, B2) have been developed (see also <http://www.grida.no/climate/ipcc/emission/index.htm>)

These data are available at the World Data Center for Climate, Hamburg.(wdc-climate.de).

Additional info

Additional information for this CERA experiment is available [here](#).

Available Datasets (Page 1 of 13)

<< Prev 25 Next 25 >>

	Name	Progress
<input type="checkbox"/>	EH5_OM_CONST_ALAKE: lake fraction of grid box	complete
<input type="checkbox"/>	EH5_OM_CONST_ALB: surface background albedo	complete
<input type="checkbox"/>	EH5_OM_CONST_AZ0L: roughness length over land	complete

Data and article

- > **The DOI system offers an easy way to connect the article with the underlying data:**

The dataset:

G.Yancheva, . R Nowaczyk et al (2007)

Rock magnetism and X-ray fluorescence spectrometry analyses on sediment cores of the Lake Huguang Maar, Southeast China, PANGAEA

[doi:10.1594/PANGAEA.587840](https://doi.org/10.1594/PANGAEA.587840)

Is cited in the article:

G. Ycheva, N. R. Nowaczyk et al (2007)

Influence of the intertropical convergence zone on the East Asian monsoon
Nature 445, 74-77

[doi:10.1038/nature05431](https://doi.org/10.1038/nature05431)

Examples of DOIs for scientific content

- > [doi:10.1594/EURORAD/CASE.1113](https://doi.org/10.1594/EURORAD/CASE.1113) in cooperation with *European Congress for Radiology (ECR)* over 6.500 Case studies
- > [doi:10.2312/EGPGV/EGPGV06/027-034](https://doi.org/10.2312/EGPGV/EGPGV06/027-034) in cooperation with *European Association for Computer Graphics (Eurographics)* over 300 articles (Grey literature)
- > [doi:10.1594/ecrystals.chem.soton.ac.uk/145](https://doi.org/10.1594/ecrystals.chem.soton.ac.uk/145) Together with the project *eBank of UK Office for Library Networking* we assigned DOI names for crystal structures.
- > [doi:10.2314/CERN-THESIS-2007-001](https://doi.org/10.2314/CERN-THESIS-2007-001) DOI names for Cern theses
- > [doi:10.2314/511535090](https://doi.org/10.2314/511535090) DOI names for final reports of projects funded by the German government
- > [doi:10.3203/iwf/C-11493](https://doi.org/10.3203/iwf/C-11493) over 12.000 DOI names for short scientific movie clips

“Non-commercial”

3 ways of financing:

- > **Customer-relation with Publication agents**
- > **Base funding of TIB (National mission)**
- > **Funding agencies (Long-time accessibility of research results)**

The new trajectory of research

