



Neue Wege des Publizierens: Die Zitierfähigkeit wissenschaftlicher Primärdaten

Jan Brase
Forschungszentrum L3S
Technische Informationsbibliothek (TIB)

**11. IUK Jahrestagung,
09.-11.05.2005
Bonn**



Projekt STD

- **Michal Diepenbroek & Uwe Schindler**
World Data Center for Marine Environmental Sciences,
MARUM Universität Bremen
- **Hannes Grobe**
Alfred-Wegener-Institut für Polar- und Meeresforschung, Bremerhaven
- **Heinke Höck & Michael Lautenschlager**
World Data Center Climate,
Max-Planck-Institut für Meteorologie Hamburg
- **Jens Klump**
GeoForschungsZentrum Potsdam
- **Irina Sens**
Technische Informationsbibliothek, Hannover



Hintergrund

Das International Council for Science (ICSU) hat 2004 in seinem Bericht "Data and information", dringend eine neue Strategie für den Umgang mit wissenschaftlichen Daten gefordert.

Der DFG-Ausschuss „*Wissenschaftliche Literaturversorgungs- und Informationssysteme*“ hat 2004 ein Projekt gestartet, um den Zugang zu wissenschaftlichen Primärdaten zu verbessern.

Die *Technische Informationsbibliothek (TIB)* soll im Rahmen dieses Projektes als zentrale Registrierungsagentur für Primärdaten fungieren. Angefangen für den Bereich der Geowissenschaften.

Das Projekt wird technisch betreut vom Forschungszentrum L3S aus Hannover, bei dem die TIB Mitglied ist.



Primärdaten

Wissenschaftlerin Frau Wetter gewinnt im Rahmen ihrer Forschung für das *Weltdatenzentrum für Klima* Primärdaten über das Wetter in Hannover im Jahre 2003.

Die Primärdaten werden wie gewohnt direkt im *Weltdatenzentrum für Klima* geprüft, evaluiert, abgespeichert und verwaltet.

Zusätzlich registriert Frau Wetter die Primärdaten auch bei der *TIB*.



- **Daten-DOI:**
 - Reine Datensätze
 - Ca. 2.000.000 jährlich
 - Nur DOI-Registrierung
- **Katalog-DOI:**
 - Datensammlungen oder Auswertungen (Sekundärdaten)
 - Ca. 200.000 jährlich
 - DOI- und URN-Registrierung
 - Zugänglich über den Bibliothekskatalog



Registrierung von Primärdaten

Frau Wetter sendet der *TIB* die URL unter der die Daten zugänglich sind und im Falle einer Katalog-DOI eine XML-Datei mit den relevanten bibliographischen Metadaten

Inklusive der Angaben, die nach ISO 690-2 zur Zitierung elektronischer Medien verlangt werden:

- Autor
- Titel
- Umfang
- Editionsnummer
- Sprache
- Verleger
- Publizierungsdatum
- Publizierungsort



Identifizierer

Die *TIB* speichert diese Informationen zu den Primärdaten ab und gibt den Primärdaten als eindeutige Identifikationsnummer einen *Digital object identifier (DOI)*.

In ihrer Publikation zitiert Frau Wetter diese Primärdaten nun durch die Angabe dieses DOI, z.B.:

**Wetter, Hilde (2003) Wetter in Hannover 2003, WDCC.
doi:10.1594 /WDCC/W_Han_2003_MMB_2**



Auflösung der DOI

Dieser DOI kann nun auf verschiedene Arten weltweit in jedem Internet Browser aufgelöst werden um Zugang zu den Daten zu erhalten:

- 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

Oder durch

- Doi://10.1594/WDCC/W_Han_2003_MMB_2

(nach Installation eines zusätzlichen Browser-Plugins)



Nutzungsszenario

Herr Sturm liest in einer Fachzeitschrift die Veröffentlichung von Frau Wetter und möchte gerne deren Daten unter anderen Gesichtspunkten auswerten.

Über die DOI hat er sofort Zugriff zu den Daten

Er besorgt sich die Erlaubnis von Frau Wetter und dem *Weltdatenzentrum für Klima* zur Verwendung der Daten..

In seiner eigenen Publikation “**Vergleich des Wetters von Hannover und Haiti**” zitiert Herr Sturm die Daten von Frau Wetter durch deren DOI und kann somit auf die Eigenständigkeit der Daten verweisen.



Uniform resource name (URN)

Durch eine Kooperation der *TIB* mit der *Deutschen Bibliothek* wird für jeden registrierten Katalog-DOI zusätzlich auch eine URN registriert. Die Struktur ist dieselbe, um eine höchstmögliche einfache Interoperabilität zu gewährleisten.

- DOI-Struktur:

10.1594/WDCC/W_Han_2003_MMB_2

- URN-Struktur:

Urn:nbn:de:tib-10.1594/WDCC/W_Han_2003_MMB_2



Technische Details



http://tws.gbv.de/CHARSET=ISO-8859-1/DB=2.63/FKT=1016/FRM=exk%2Bprimaerdaten/IMPLAND=Y/LNG=DU/LRSET=4/SET=4/SID=17b67f59-9d/SRT=YOP/TTL=1/SHW?FRST=3

New Tab TIBORDER - Dokumentlieferdienst d...

Suchen | **Suchergebnis** | Erweiterte Suche | Zwischenablage | Hilfe

© 1998-2005 OCLC PICA

TIBORDER

Katalog

suchen [und] sortiert nach

suchen

Kundennummer: | [Datenbanken](#)

Suchgeschichte | **Kurzliste** | Titeldaten

Download
Zwischenablage

■ Ihre Aktion suchen [und] exk primaerdaten

3 von 29

Datenbanken
Bestellung ohne
Recherche
Benutzerinfo
TIB Homepage

Titel: [IPCC-DDC_CSIRO_SRES_B1: 140 YEARS MONTHLY MEANS Commonwealth Scientific and Industrial Research Organisation Australia](#) / World Data Center for Climate (WDCC) , Hamburg .Hal@Gordon

Beteiligt: [Hal Gordon](#)

Körperschaft: [World Data Center](#) for [Climate](#) (WDCC)

Erschienen: 2005-02-20

Umfang: Online-Ressource (280217280 Bytes).

Anmerkung: Mode: Abstract
StructuralType: Digital
CreationDate: 2001-12-14

Inhalt: The SRES data sets were published by the IPCC in 2000 and classified into four different scenario families (A1, A2, B1, B2). SRES_B1 storyline describes a global population similar to A1 but with rapid change in economic structure toward a service and information economy, with reduction in material intensity and the introduction of clean and resource efficient technologies.
The Mark 2 CSIRO coupled global climate model consists of the atmospheric component which has 9 vertical levels with a R21 horizontal resolution and the ocean component which was based on the GFDL code.
M2CSIRO (http://www.dar.csiro.au/publications/hennessy_1998a.html#ccm) The changes of anthropogenic emissions of CO2, CH4, N2O and sulphur dioxide are prescribed according to the above mentioned scenario.

Technische Angaben: Format: GRIB
Volltext/Image: [Gesamtes Dokument](#) lizenzfrei!
[URN](#) lizenzfrei!

Bestandsinfo: [Anzeigen](#) lizenzfrei!
Anmerkung: **Primaerdaten**

3 von 29



Resolving

Die DOI werden über das Handle System von der CNRI (Cooperation for National Research Initiatives) aufgelöst.

“A free java based comprehensive system for assigning, managing, and resolving persistent identifiers, known as "handles," for digital objects and other resources on the Internet”

Verwendet vom IDF, aber auch von der MIT Library, Library of Congress,



Bibliographische Metadaten für Primärdaten

Attribut	LOM (DC) mapping
1. DOI	dc:identifier
2. identifier	dc:identifier
3. creator	dc:creator
4. publisher	dc:publisher
5. title	dc:title
6. language	dc:language
7. StructuralType (<i>digital</i>)	dc:format
8. mode (<i>abstract</i>)	none
9. resourceType (<i>dataset</i>)	dc:format



Attribute	LOM (DC) mapping
10. registrationAgency (<i>10.1594</i>)	dc:contributor
11. issueDate	dcterms:issued
12. issueNumber	dc:identifier
13. creationDate	dcterms:created
14. publicationDate	dc:date
15. description	dc:description
16. publicationPlace	none
17. size	dcterms:extend
18. format	dc:format
19. edition	none
20. relatedDOIs	dc:source (and others)



Qualität der Metadaten

- Es ist offensichtlich, dass zusätzliche Metadaten über den Inhalt die Qualität der Registrierung steigern werden.
- Es ist allerdings möglich Volltextsuchen über das Attribut Description auszuführen.
- In der ersten Phase des Projekts war die Zitierfähigkeit der Primärdaten zusammen mit der Registrierung der bibliographischen Metadaten (inklusive eines guten Mappings zum Standard Dublin Core) das Hauptziel. In späteren Phasen wird aber über eine Erweiterung der Metadaten nachgedacht werden.



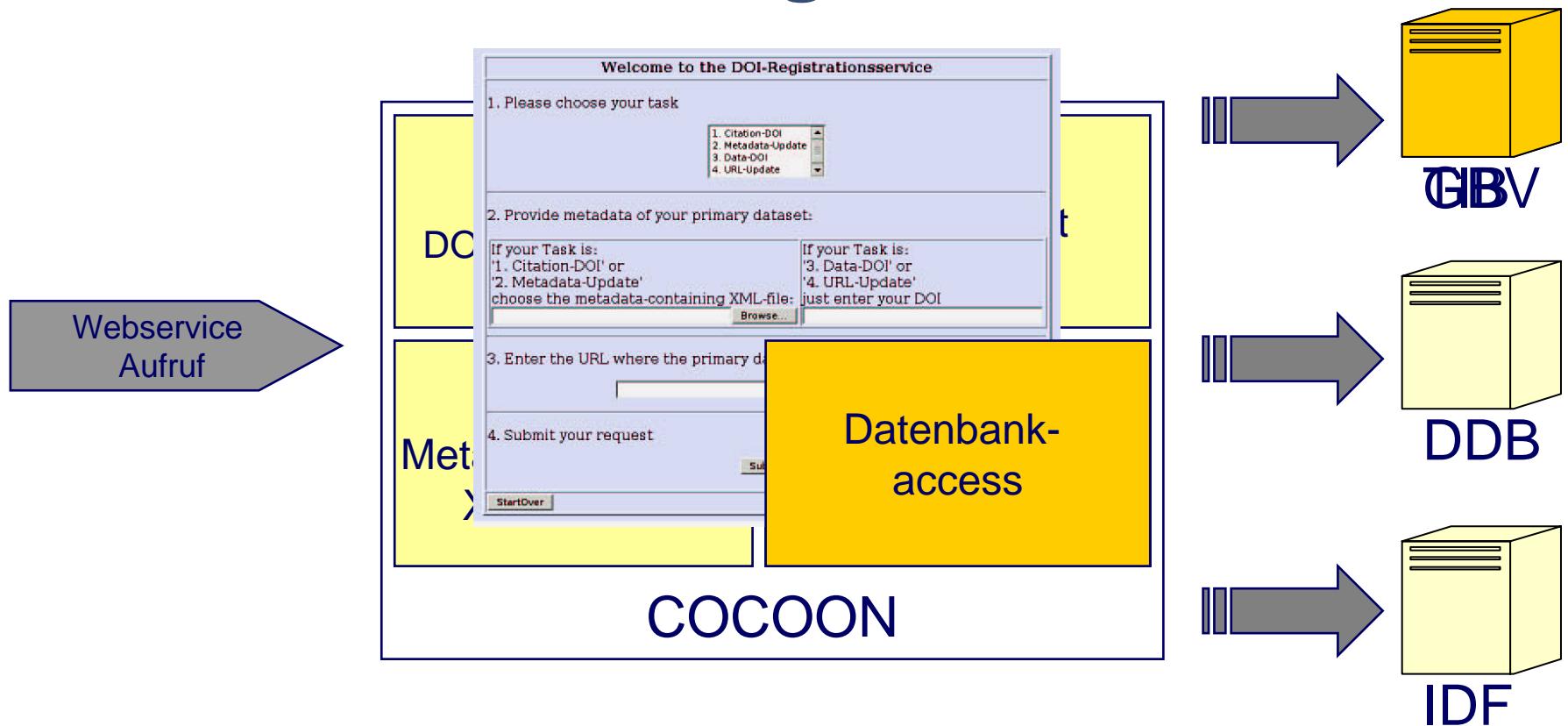
Einschätzung

- Bisher gibt es keine eigenständigen Publikation von Primärdaten.
- In Fachzeitschriften ist bisher kein Platz für Neuauswertungen bestehender Daten.
- Projektdaten sind häufig schlecht dokumentiert und daher schlecht zugänglich.
- Große Mengen von Primärdaten können nicht verwendet werden, da sie nicht verfügbar und nur kleinen Kreisen bekannt sind.
- Die Möglichkeit der Zitierbarkeit von Primärdaten kann diese Situation fundamental ändern.

 In der Wissenschaftlichen Welt besteht ein großes Interesse an der Registrierung von Primärdaten.



Technische Realisierung in der TIB





Stand des Projekts

Für den Bereich der Geowissenschaft fungiert die *TIB* in Kooperation mit

- World Data Center climate & mare
- GeoForschungsZentrum Potsdam
- Alfred Wegener Institut

als weltweit erste zentrale DOI Registrierungsagentur für Primärdaten.

Zum Stand 05/2005 sind die ersten 175.000 Datensätze in der *TIB* registriert.

Zusätzlich sind schon 30 Datensätze über den Katalog der *TIB* zugänglich.



Zukunft

2006 wird die Registrierung auf andere Fachbereiche ausgeweitet werden.

Ab 2007 wird die Registrierung **jeder** Art von naturwissenschaftlichen Inhalten Kernaufgabe der TIB.

Dieses könnte umfassen:

- Bodenproben
- 3D-Modelle
- Graue Literatur
- ...



Mögliche Strukturen

Wir fangen gerade erst an!

Gesucht werden Publikations-Agenten, die die Inhalte sammeln, evaluieren und warten.

Projekten könnte Geld zur Verfügung gestellt werden zur Registrierung der Inhalte (in Verbindung mit der Nachweispflicht vom BMBF).

Die Registrierung von wiss. Inhalten muss **nicht-kommerziell** sein und bleiben.



Weitere Informationen

Projekt Webseite:

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

TIB Handle Server:

<http://doi.tib-hannover.de:8000>

DOI Foundation:

<http://www.doi.org>

URN Registrierung der DDB:

<http://www.persistent-identifier.de>



Vielen Dank! Noch Fragen?

