DataCite – Bridging the gap and helping to find, access and reuse data

Herbert Gruttemeier

OpenAIREplus workshop
February 8th, 2013
Braga

Nucleic Acids Research

Table of Contents

Volume 33 Issue 15

Clear Get All Checked Abstracts

Methods Online

Kei Moriyama, Michiko Kimoto, Tsuneo Mitsui, Shigeyuki Yokoyama, and Ichiro Hirao
Site-specific biotinylation of RNA molecules by transcription using unnatural base pairs
Nucl. Acids Res. 33(15): e129 doi:10.1093/nar/gni128
» Abstract » FREE Full Text (HTML) » Full Text (PDF) » Screen PDF Supplementary Material
OPEN ACCESS (cc) BY-NC

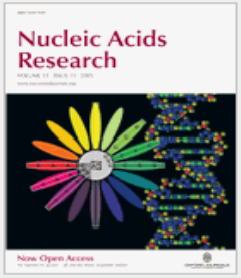
Tom Ebersole, Yasuhide Okamoto, Vladimir N. Noskov, Natalay Kouprina, Jung-Hyun Kim, Sun-Hee Leem, J. Carl Barrett, Hiroshi Masumoto, and Vladimir Larionov
Rapid generation of long synthetic tandem repeats and its application for analysis in human artificial chromosome formation
Nucl. Acids Res. 33(15): e130 doi:10.1093/nar/gni129
» Abstract » FREE Full Text (HTML) » Full Text (PDF) » Screen PDF Supplementary Material
OPEN ACCESS (cc) BY-NC

Amarsanaa Jazag, Fumihiko Kanai, Hideaki Ijichi, Keisuke Tateishi, Tsuneo Ikenoue, Yasuo Tanaka, Miki Ohta, Jun Imamura, Bayasi Guleng, Yoshinari Asaoka, Takao Kawabe, Makoto Miyagishi, Kazunari Taira, and Masao Omata
Single small-interfering RNA expression vector for silencing multiple transforming growth factor-β pathway components
Nucl. Acids Res. 33(15): e131 doi:10.1093/nar/gni130
» Abstract » FREE Full Text (HTML) » Full Text (PDF) » Screen PDF
OPEN ACCESS (cc) BY-NC

Ding-Gan Liu and Li Sun
Direct isolation of specific RNA-interacting proteins using a novel affinity medium
Nucl. Acids Res. 33(15): e132 doi:10.1093/nar/gni133
» Abstract » FREE Full Text (HTML) » Full Text (PDF) » Screen PDF
OPEN ACCESS (cc) BY-NC

« Previous | Next Issue »

This Issue
33 (15)



» Index By Author
» Cover Image

› Methods Online
› Articles
› Corrigendum

Find articles in this issue containing these words:
Search This Issue GO

Advance Access

Search this journal:
GO
Advanced »

Current Issue
March 2012 40 (6)

Alert me to new issues

The Journal

About this journal
NAR Methods online
2012 Database Issue
2011 Web Server Issue
NAR Special Collections
Referee Information
Rights & Permissions
Dispatch date of the next issue
This journal is a member of the Committee on Publication Ethics (COPE)
view Recent Comments on articles
We are mobile – find out more

Impact factor: 7.836
5-Yr impact factor: 7.314

56. Kruger, D.H., Kupper, D., Meisel, A., Reuter, M., Schroeder, C. (1995) The significance of distance and orientation of restriction endonuclease recognition sites in viral DNA genomes *FEMS Microbiol. Rev.*, **17**, 177–184[CrossRef][ISI][Medline] .

57. Saha, S. and Rao, D.N. (1997) Mutations in the Res subunit of the EcoPI restriction enzyme that affect ATP-dependent reactions *J. Mol. Biol.*, **269**, 342–354[CrossRef][ISI][Medline] .

Science 17 January 1997:
Vol. 275. no. 5298, pp. 377 - 380
DOI: 10.1126/science.275.5298.377

REPORTS

Kinetic Measurement of the Step Size of DNA Unwinding by *Escherichia coli* UvrD Helicase

Janid A. Ali and Timothy M. Lohman *

The kinetic mechanism by which the DNA repair helicase UvrD of *Escherichia coli* unwinds duplex DNA was examined with the use of a series of oligodeoxynucleotides with duplex regions ranging from 10 to 40 base pairs.

65. Francino, M.P. and Ochman, H. (1997) Strand asymmetries in DNA evolution *Trends Genet.*, **13**, 240–245[CrossRef][ISI][Medline] .

66. Frank, A.C. and Lobry, J.R. (1999) Asymmetric substitution patterns: a review of possible underlying mutational or selective mechanisms *Gene*, **238**, 65–77[CrossRef][ISI][Medline] .

67. Beletskii, A., Grigoriev, A., Joyce, S., Bhagwat, A.S. (2000) Mutations induced by bacteriophage T7 RNA polymerase and their effects on the composition of the T7 genome *J. Mol. Biol.*, **300**, 1057–1065[CrossRef][ISI][Medline] .

68. Studier, F.W. (1975) Gene 0.3 of bacteriophage T7 acts to overcome the DNA restriction system of the host *J. Mol. Biol*, **94**, 283–295[CrossRef][ISI][Medline] .

Publishers' data policies

INTERNATIONAL ASSOCIATION OF SCIENTIFIC, TECHNICAL & MEDICAL PUBLISHERS

The voice of research publishing for 40 years

stm

Events | Document Library | Copyright & Legal Affairs | Standards & Technology | Public Affairs | Research4Life Outreach Programmes

You are here: Home > Public Affairs > Introduction

Brussels Declaration on STM Publishing

by the international scientific, technical and medical (STM) publishing community as represented by the individual publishing houses and publishing trade associations, who have indicated their assent below.

You can download this as a PDF document.

Many declarations have been made about the need for particular business models in the STM information community. STM publishers have largely remained silent on these matters as the majority are agnostic about business models: what works, works. However, despite very significant investment and a massive rise in access to scientific information, our community continues to be beset by propositions and manifestos on the practice of scholarly publishing. Unfortunately the measures proposed have largely not been investigated or tested in any evidence-based manner that would pass rigorous peer review. In the light of this, and based on over ten years experience in the economics of online publishing and our longstanding collaboration with researchers and librarians, we have decided to publish a declaration of principles which we believe to be self-evident.

1. **The mission of publishers is to maximise the dissemination of knowledge through economically self-sustaining business models.** We are committed to change and innovation that will make science more effective. We support academic freedom: authors should be free to choose where they publish in a healthy, undistorted free market
2. **Publishers organise, manage and financially support the peer review processes of STM journals.** The imprimatur that peer-reviewed journals give to accepted articles (registration, certification, dissemination and editorial improvement) is irreplaceable and fundamental to scholarship
3. **Publishers launch, sustain, promote and develop journals for the benefit of the scholarly community**
4. **Current publisher licensing models are delivering massive rises in scholarly access to research outputs.** Publishers have invested heavily to meet the challenges of digitisation and the annual 3% volume growth of the international scholarly literature, yet less than 1% of total R&D is spent on journals
5. **Copyright protects the investment of both authors and publishers.** Respect for copyright encourages the flow of information and rewards creators and entrepreneurs
6. **Publishers support the creation of rights-protected archives that preserve scholarship in perpetuity**
7. **Raw research data should be made freely available to all researchers.** Publishers encourage the public posting of the raw data outputs of research. Sets or sub-sets of data that are submitted with a paper to a journal should wherever possible be made freely accessible to other scholars
8. **Publishing in all media has associated costs.** Electronic publishing has costs not found in print publishing. The costs to deliver both are higher than print or electronic only. Publishing costs are the same whether funded by supply-side or demand-side models. If readers or their agents (libraries) don't fund publishing, then someone else (e.g. funding bodies, government) must

Public Affairs Menu

- Introduction
- Brussels Declaration
- PEER
- Press Releases & Announcements
- Public Affairs Committee

lineage-minus (Lin-), the antibodies and fluorochromes that are contained in the 'cocktail' need to be specified for the 'dump' channel.

Top of page

Sharing data sets

A condition of publication in a Nature journal is that authors are required to make materials, data and associated protocols promptly available to others without preconditions.

Data sets must be made freely available to readers from the date of publication, and must be provided to editors and peer-reviewers at submission, for the purposes of evaluating the manuscript.

For the following types of data set, submission to a community-endorsed, public repository is mandatory. Accession numbers must be provided in the paper. Examples of appropriate public repositories are listed below.

DNA and protein sequences

Protein sequences: Uniprot

DNA and RNA sequences: Genbank/EMBL/DDBJ, Protein DataBank, UniProt.

DNA sequencing data (traces for capillary electrophoresis and short reads for next-generation sequencing): NCBI trace and short-read archive, EBI Ensembl trace sever

Deep sequencing data: deposit in GEO or ArrayExpress upon submission to the journal. Accession numbers must be provided in the published manuscript.

This policy includes even short stretches of novel sequence information such as **epitopes, functional domains, genetic markers, or haplotypes**. Short novel sequences must include surrounding sequence information to provide context.

The sequences of all **RNAi, antisense and morpholino probes** must be included in the paper or deposited in a public database, with the accession number quoted. When an unpublished library is included in the paper, at minimum the sequences of the probes central to the conclusions of the paper must be presented.

Macromolecular structures

Authors of papers describing structures of biological macromolecules must provide atomic coordinates and related experimental data (structure factor amplitudes/intensities for crystal structures, or restraints for NMR structures) upon request of editors for the purposes of evaluating the manuscript, if they are not already freely accessible in a publicly available and recognized database (for example, Protein DataBank, Uniprot, Nucleic Acids Database or Biological Magnetic Resonance Databank).

Publishers' data policies

extract from

Nature Publishing Group, Editorial Policies, Availability of data and materials

Geochimica et Cosmochi
Journal of The Geochemical Society

ISSN: 0016-7037
Imprint: ELSEVIER

(1989b)), then two-author papers alphabetically (Black and Brown (1991), Black and Brown (1992), Black and Greene (1987)), then three or more authors chronologically (Black, Brown and Blue (1989), Black, Blue and Brown (1991)). Authors should take care that all literature citations, in figure captions and tables as well as main text, have accompanying entries in the References, and also that there are no superfluous entries.

Supplementary data

Elsevier accepts electronic supplementary material to support and enhance your scientific research. Supplementary files offer the author additional possibilities to publish supporting applications, high-resolution images, background datasets, sound clips and more. Supplementary files supplied will be published online alongside the electronic version of your article in Elsevier Web products, including ScienceDirect: ⊟→ http://www.sciencedirect.com. In order to ensure that your submitted material is directly usable, please provide the data in one of our recommended file formats. Authors should submit the material in electronic format together with the article and supply a concise and descriptive caption for each file. For more detailed instructions please visit our artwork instruction pages at http://www.elsevier.com/artworkinstructions.

NOTE: Supplementary materials published online only in GCA are referred to as Electronic Annexes. Use of an Electronic Annex is required for extensive supplementary materials such as large data tables, sample descriptions, multiple figures, etc that would otherwise occupy a large amount of printed space.

Data at PANGAEA

Electronic archiving of supplementary data enables readers to replicate, verify and build upon the conclusions published in your paper. We recommend that data should be deposited in the data library PANGAEA (⊟→ http://www.pangaea.de). Data are quality controlled and archived by an editor in standard machine-readable formats and are available via Open Access. After processing, the author receives an identifier (DOI) linking to the supplements for checking. As your data sets will be citable you might want to refer to them in your article. In any case, data supplements and the article will be automatically linked as in the following example: ⊟→ doi:10.1016/0016-7037(95)00105-9. Please use PANGAEA's web interface to submit your data (⊟→ http://www.pangaea.de/submit/).

Google Maps and KML files

KML (Keyhole Markup Language) files (optional): You can enrich your online articles by providing KML files which will be visualized using Google maps. The KML files can be uploaded in our online submission system. KML is an XML schema for expressing geographic annotation and visualization within Internet-based Earth browsers. Elsevier will generate Google Maps from the submitted KML files and include these in the article when published online. Submitted KML files will also be available for downloading from your online article on ScienceDirect. For more information see http://www.elsevier.com/googlemaps.

After Acceptance

Use of the Digital Object Identifier

The Digital Object Identifier (DOI) may be used to cite and link to electronic documents. The DOI consists of a unique alpha-numeric character string which is assigned to a document by the publisher upon the initial electronic publication. The assigned DOI never changes. Therefore, it is an ideal medium for citing a document, particularly 'Articles in press' because they have not yet received their full bibliographic information. The correct format for citing a DOI is shown as follows (example taken from a document in the journal *Physics Letters B*):
doi:10.1016/j.physletb.2010.09.059
When you use the DOI to create URL hyperlinks to documents on the web, the DOIs are guaranteed never to change.

Proofs

One set of page proofs (as PDF files) will be sent by e-mail to the corresponding author (if we do not have an e-mail address then paper proofs will be sent by post) or, a link will be provided in the e-mail so that authors can download the files themselves. Elsevier now provides authors with PDF proofs which can be annotated; for this you will need to download Adobe Reader version 7 (or higher) available free from ⊟→ http://get.adobe.com/reader. Instructions on how to annotate PDF files will accompany the proofs (also given online). The exact system requirements are given at the Adobe site: ⊟→ http://www.adobe.com/products/reader/tech-specs.html.
If you do not wish to use the PDF annotations function, you may list the corrections (including replies to the Query Form) and return them to

Geochimica et Cosmochimica Acta

Volume 59, Issue 11, June 1995, Pages 2265–2278

Article

Reconstructing past sea surface temperatures: Correcting for diagenesis of bulk marine carbonate

Daniel P. Schrag[1, 2, *], Donald J. DePaolo[1, 2], Frank M. Richter[3]

[1] Berkeley Center for Isotope Geochemistry, Department of Geology and Geophysics, University of California at Berkeley, Berkeley, CA 94720, USA

[2] Earth Science Division, Lawrence Berkeley Laboratory, Berkeley, CA 94720, USA

[3] Department of the Geophysical Sciences, University of Chicago, Chicago, IL 60637, USA

Received 4 December 1993. Accepted 3 March 1995. Available online 5 April 2000.

http://dx.doi.org/10.1016/0016-7037(95)00105-9, How to Cite or Link Using DOI

nt and profile measurements

Site Search

GO

About the DOI System
Overviews & Standards
DOI & URI
Frequently Asked Questions
Factsheets
DOI® Handbook

International DOI Foundation
Welcome Message
Membership

Activities
News/Events
Mailing Lists/Working Groups
Reviews

Resources
Registration Agencies
White Papers
Examples
Tools
DOI System Proxy Server
Vocabulary Mapping Framework

IDF Members Only
Member's Site
RAWG Site

Subscribe to Free DOI® News!

Home

Welcome to the DOI® System

The Digital Object Identifier (DOI®) System is for identifying content objects in the digital environment. DOI® names are assigned to any entity for use on digital networks. They are used to provide current information, including where they (or information about them) can be found on the Internet. Information about a digital object may change over time, including where to find it, but its DOI name will not change.

The DOI System provides a framework for persistent identification, managing intellectual content, managing metadata, linking customers with content suppliers, facilitating electronic commerce, and enabling automated management of media. DOI names can be used for any form of management of any data, whether commercial or non-commercial. The DOI System is an ISO International Standard.

The system is managed by the International DOI Foundation, an open membership consortium including both commercial and non-commercial partners. Over 55 million DOI names have been assigned by DOI System Registration Agencies in the US, Australasia, and Europe.

Using DOI names as identifiers makes managing intellectual property in a networked environment much easier and more convenient, and allows the construction of automated services and transactions.

To learn more about DOI names, see the Overviews, and begin with the Introductory Overview and Introductory Slide Presentation. The Factsheets, including "Managing Data Relationships Using DOI Resolution" and "DOI System and Standard Identifier Schemes" discuss key topics about the system. For the most complete description of all aspects of DOI System technology and policy, consult the DOI® Handbook.

In the News

Japan Link Center Appointed as New Registration Agency in Japan

IDF expands Chinese Registration Agency activities

Entertainment ID Registry (EIDR) Gains Strong Support from Major Entertainment Companies

Airiti Inc. Appointed as New DOI Registration Agency

DOI System approved as an International Standard

DOI System exceeds 50 million assigned identifiers

DOI System and Linked Data

Revised DOI Charging model

[Earlier News]

Resolve a DOI Name

Type or paste a DOI name (e.g., 10.1000/182) into the text box below.

Submit

Updated 21 March 2012

http://www.doi.org

Handle System®

Unique and Persistent Identifiers for Internet Resources

HOME | ABOUT THE HANDLE SYSTEM | DOCUMENTATION | SOFTWARE | REGISTRATION | TECHNICAL SUPPORT

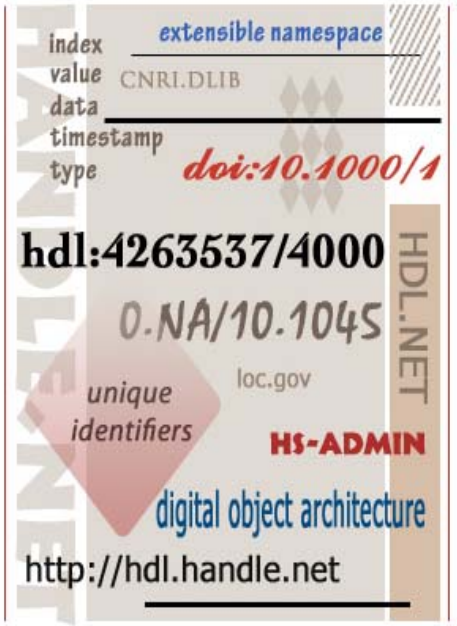
Welcome to the Handle System

Current Release – Version 7.0

The Handle System provides efficient, extensible, and secure resolution services for unique and persistent identifiers of digital objects, and is a component of CNRI's Digital Object Architecture. Digital Object Architecture provides a means of managing digital information in a network environment. A digital object has a machine and platform independent structure that allows it to be identified, accessed and protected, as appropriate. A digital object may incorporate not only informational elements, i.e., a digitized version of a paper, movie or sound recording, but also the unique identifier of the digital object and other metadata about the digital object. The metadata may include restrictions on access to digital objects, notices of ownership, and identifiers for licensing agreements, if appropriate.

• • • •

The Handle System includes an open set of protocols, a namespace, and a reference implementation of the protocols. The protocols enable a distributed computer system to store identifiers, known as handles, of arbitrary resources and resolve those handles into the information necessary to locate, access, contact, authenticate, or otherwise make use of the resources. This information can be changed as needed to reflect the current state of the identified resource without changing its identifier, thus allowing the name of the item to persist over changes of location and other related state information. Some examples of applications that use HDL® identifier and resolution services as infrastructure are rights management applications, digital object registries and repositories, and institutional data preservation and archiving.



Corporation for National Research Initiatives (CNRI) • 1895 Preston White Drive • Reston, Virginia • 20191 • 703.620.8990 • July 2011

At the infrastructure level, DOI names are handles.

Get the hdl enabled Logo • Privacy Policy

http://www.handle.net

From KE workshop presentation, The Hague, June 2011 (L. Lannom)

Handle System Usage

- Library of Congress
- DTIC (Defense Technical Information Center)
- IDF (International DOI Foundation)
  - CrossRef (scholarly journal consortium, representing >2K publishers & societies)
  - DataCite (consortium of 19 members from 11 countries, started by TIB)
  - EIDR (Entertainment Identifier Registry)
  - mEDRA (Multilingual European DOI Registration Agency)
  - R.R. Bowker (bibliographic data - ISBN)
  - Office of Publications of the European Community (OPOCE)
  - Wanfang Data
- OECD
- National Agricultural Library/USDA
- DSpace (MIT + HP)
- ADL (DoD Advanced Distributed Learning initiative)
- Australian National Data Service (ANDS)
- EPIC (European Persistent Identifier Consortium)
- GENI (Global Environment for Network Innovations)

From KE workshop presentation, The Hague, June 2011 (L. Lannom)

Q4 – Policies: Does your organization have a policy on (the quality of) Persistent Identifiers?

- Again, the Handle System is a tool for building infrastructure. Use is not actively monitored. That said,
  - CNRI controls the distribution of prefixes
  - Prefixes handed out one at a time or in batches to known entities, e.g., IDF
  - Small fee charged for production use to discourage abuse
  - Prefix holders must agree to be good citizens

From KE workshop presentation, The Hague, June 2011 (N. Paskin)

Policies

doi>

Policies: Does your organisation have a policy on (the quality of) Persistent Identifiers? What are or should be the main elements in it?

- Yes: *raison d'etre* of organisation
- Conformance to standard
- Conformance to policy (proxies, display, trademark, etc.)
- Agreement to share funding
- Legal agreement for membership of IDF, with obligations
- Governance structure of IDF
- Persistence of DOI resolution:
  - Change of management (registrant); change/cessation of RA
  - Already tested
- Persistence of underlying technologies:
  - CNRI Handle
  - VMF/ COA

Digital Object Identifier

Pour les articles homonymes, voir DOI.

plutôt: *identifiant numérique d'objet*

Digital object identifier (DOI, littéralement « identifiant d'objet numérique ») est un mécanisme d'identification de ressources numériques, comme un film, un rapport, des articles scientifiques, etc. C'est une alternative aux URI. Depuis février 2010, l'Institut de l'information scientifique et technique (INIST, du CNRS), est doté d'un statut « agence DOI », faisant partie du consortium DataCite. Les DOI permettent notamment de faciliter l'utilisation des bases de données bibliographiques ainsi que les logiciels de gestion bibliographique.

Sommaire [masquer]

« The objects identified by DOI names may be of any form - digital, physical, or abstract - as all these forms may be necessary parts of a content management system. The DOI system is an abstract framework which does not specify a particular context of its application, but is designed with the aim of working over the Internet. »

Norman Paskin, « Digital Object Identifier (DOI®) System »

Description [modifier]

Un DOI est un cas particulier d'ider

La motivation principale pour tenter
et la motivation principale pour tent
système de l'ICANN et des registre

Le protocole de résolution, concurr
quasiment jamais utilisé Handle (q
renvoie automatiquement vers l'artic

Par exemple `hdl:cnri.dlib/dece`
`10.1340/309registries`).

`10.xxxx` est le préfixe. Il identifie le registre ou *Naming Authority*. Tout ce qui est après la barre oblique / dépend du registre. DOI a donc une infrastructure sociale (registres et bureaux d'enregistrement - *Registration Agencies*) propre. Le but est, par exemple, d'assurer la persistance des identificateurs.

Au bout d'un DOI, on trouve :

- les métadonnées (restrictions d'usage ou bien droit d'auteur, par exemple), décrites par un modèle de données commun à tous les DOI, le *indecs Data Dictionary*,
- une adresse ou localisation physique (en général un URL), le traducteur cité plus haut redirige vers cet URL,
- diverses informations, comme l'autorité de nommage.

Notes et références [modifier]

1. ↑ (en) S. Sun, S. Reilly, L. Lannom, J. Petrone, « Request for Comments 3652: Handle System Protocol (ver 2.1) Specification [archive] », novembre 2003, 52 p.

DataCite

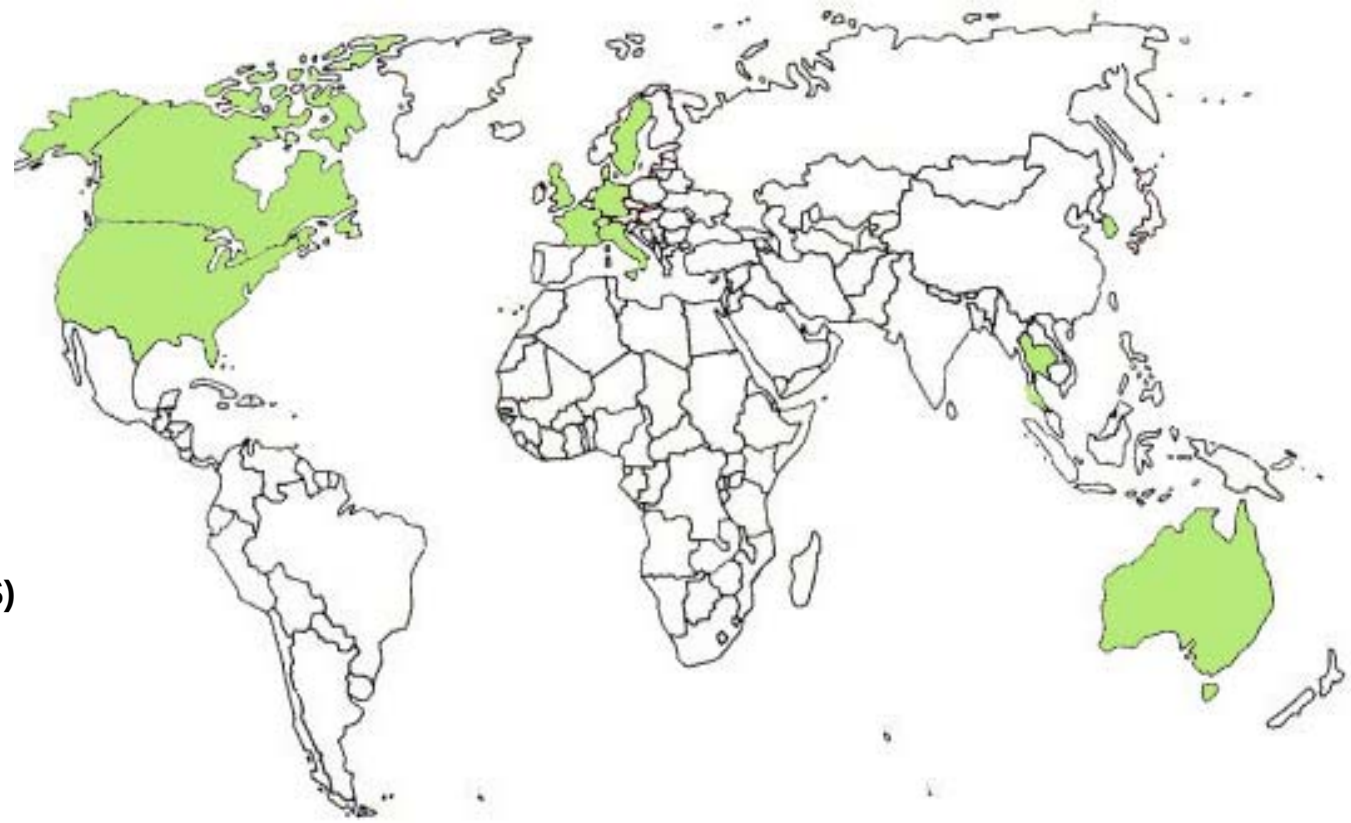
- Global consortium carried by local institutions
- Focused on improving the scholarly infrastructure around datasets and other non-textual information
- Focused on working with data centres and organisations that hold data
- Providing standards, workflows and best-practice
- Initially, but not exclusively based on the DOI system

- Memorandum of Understanding, Paris, February 2009
- Officially founded December 1st 2009 in London

DataCite

DataCite Members

- Technische Informationsbibliothek (TIB), Germany
- Canada Institute for Scientific and Technical Information (CISTI)
- California Digital Library, USA
- Purdue University, USA
- Office of Scientific and Technical Information (OSTI), USA
- The British Library
- Technical Information Center of Denmark (DTU)
- Library of TU Delft, The Netherlands
- ZBMed, Germany
- ZBW, Germany
- GESIS, Germany
- Library of ETH Zürich, Switzerland
- Institut de l'Information Scientifique et Technique (INIST-CNRS), France
- Swedish National Data Service (SND)
- Australian National Data Service (ANDS)
- Conferenza dei Rettori delle Università Italiane (CRUI)
- National Research Council of Thailand (NRCT)



Affiliated members:

- Digital Curation Center, UK
- Microsoft Research
- Interuniversity Consortium for Political and Social Research (ICPSR), USA
- Korea Institute of Science and Technology Information (KISTI)
- Bejiing Genomic Institute (BGI)

DataCite

DataCite

The DataCite registration agency

- Maintains the resolution infrastructure
- Maintains a searchable database of metadata
- Manages the identifiers over the long term
- Establishes and shares best practice

Publishing agents (data centres, research institutes, data publishers) are responsible for

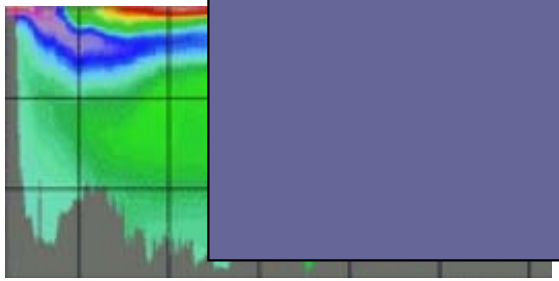
- Quality assurance
- Content storage and access
- Creating the identifiers
- Creating and updating metadata

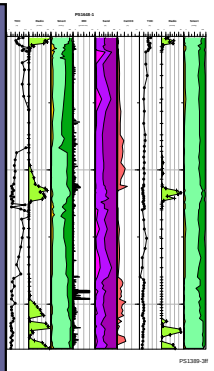
What type of data are we talking about?

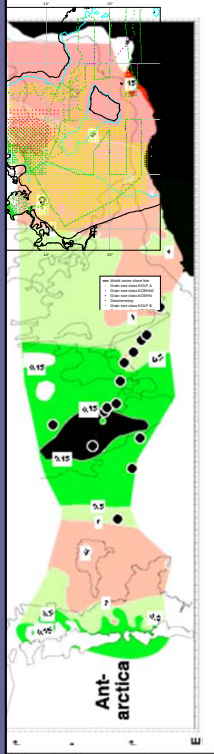
- Earth qu
  doi:10.1
- Climate
- Sea bed
- Distribu
- Medical
  270407
- Comput
- Audio re
- Grey Lit
- Videos :

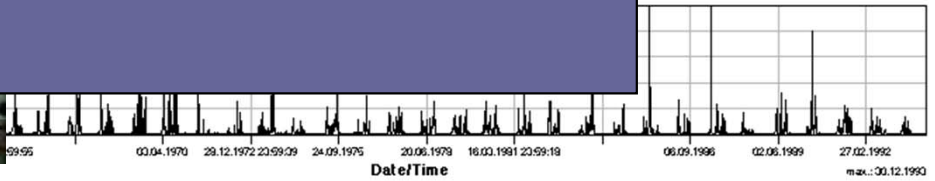
Anything that is the foundation of further research is research data

Data is evidence

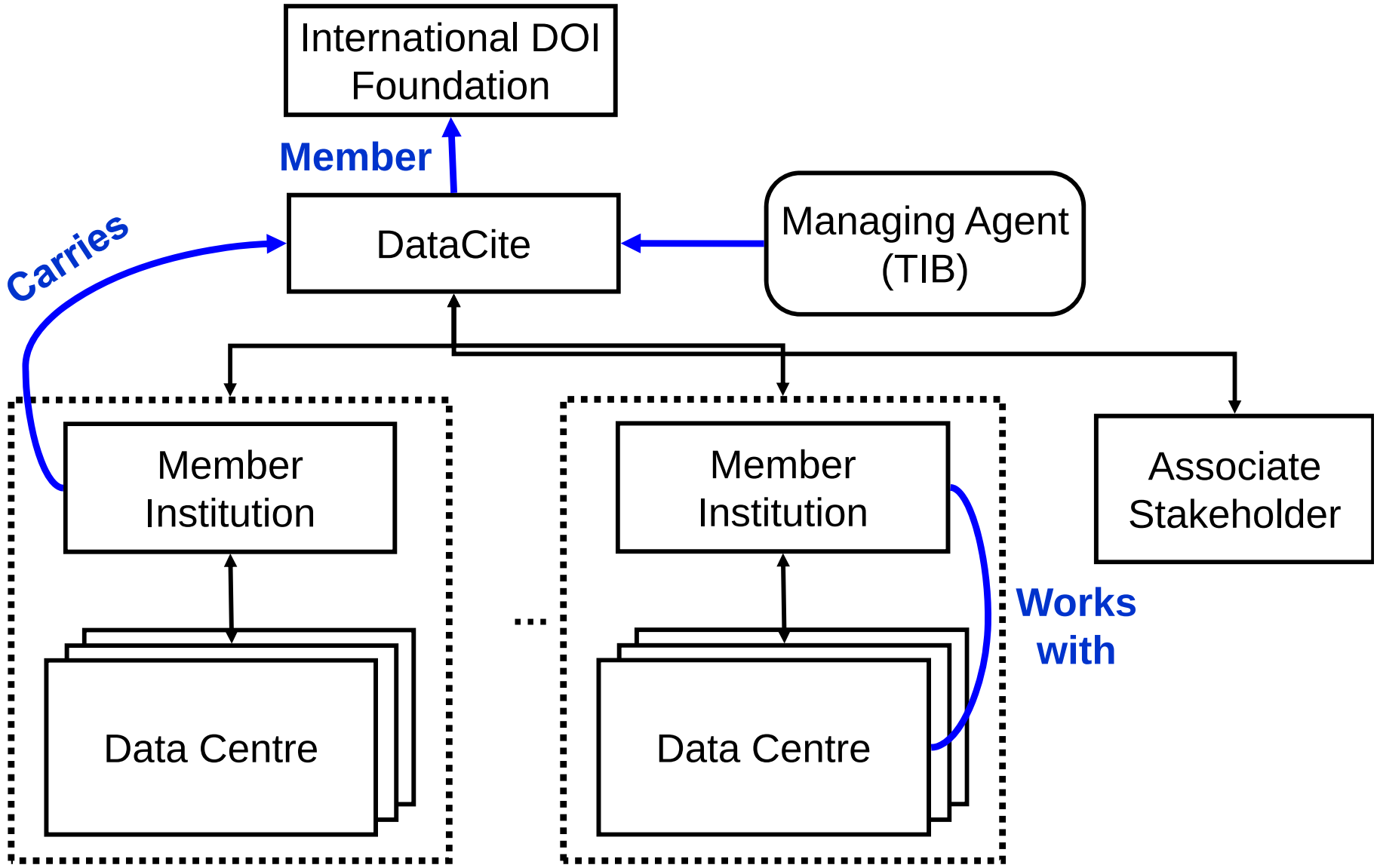








DataCite Structure



Bridging the gap

DOIs in Use: DataCite

CrossRef has registered more than 51 million DOIs on behalf of scholarly publishers. But CrossRef DOIs are not the only DOIs available in the scholarly community. DOIs for datasets associated with scholarly research are being registered by institutions in the DataCite network. DataCite and CrossRef have committed to the interoperability of their DOIs. Ideally, scholarly content like journals will cite related data by the appropriate DataCite DOI, and in return, the data record will cite the relevant article's CrossRef DOI.

(from CrossRef Quarterly, January 2012)

DataCite

Helping you to find, access, and reuse data

Why cite data?

Joint statement from STM and DataCite

Published by Jan Brase on 14 June 2012 - 1:02pm

During the DataCite su
Copenhagen, DataCite
Association today sign
joint statement to en
and data centers to li
underlying data:

. To improve the availa
of research data, Dat
encourage authors of
deposit researcher validated data in trustworthy an
Archives.

. DataCite and STM encourage Data Archives to enab
linking between datasets and publications by using
community endorsed unique persistent identifiers su
accession codes and DOI names.

. DataCite and STM encourage publishers to make vis
visibility of these links from publications to datasets

. DataCite and STM encourage Data Archives to make
visibility of these links from datasets to publications

. DataCite and STM support the principle of data re-u
purpose actively
participate in initiatives for best practice recommen
citation of datasets.

crossref

August 10, 2012

CrossRef Joins STM-DataCite Statement

In June 2012, DataCite and the International Association of STM Publishers (STM) issued a joint statement on the Linkability and Citability of Research Data (http://www.stm-assoc.org/2012_06_14_STM_DataCite_Joint_Statement.pdf). CrossRef is pleased to join and support this statement and the best practices for data it recommends.

CrossRef, a not-for-profit association of representing 4,000 scholarly publisher with 55 million content items (journal and conference proceeding articles and books and book chapters), is committed to the interoperability of CrossRef and DataCite's services which are based on the Digital Object Identifier (DOI) System, recently approved as an ISO Standard (**ISO 26324:2012, *Information and documentation - - Digital object identifier system***).

Specifically, CrossRef encourages publishers to use DataCite DOIs to link to data sets referenced in the published literature, and encourages authors of research papers to use CrossRef DOIs to link from data deposited in DataCite repositories to the published articles that draw on that data. CrossRef and DataCite are also collaborating on joint services, such as DOI Content Negotiation (http://crosscite.org/cn/), to enable publishers and data repositories to automatically interlink their content.

Data citation

Connecting article and underlying data via DOI:

The dataset:

Storz, D et al. (2009):

Planktic foraminiferal flux and faunal composition of sediment trap L1_K276 in the northeastern Atlantic.

http://dx.doi.org/10.1594/PANGAEA.724325

Is supplement to the article:

Storz, David; Schulz, Hartmut; Waniek, Joanna J; Schulz-Bull, Detlef; Kucera, Michal (2009): *Seasonal and interannual variability of the planktic foraminiferal flux in the vicinity of the Azores Current.*

Deep-Sea Research Part I-Oceanographic Research Papers, **56(1),** 107-124,

http://dx.doi.org/10.1016/j.dsr.2008.08.009

PANGAEA®
Publishing Network for Geoscientific & Environmental Data

Data Description

RIS

Citation: Storz, D et al. (2009): Planktic foraminiferal flux and faunal composition of sediment trap L1_K276 in the northeastern Atlantic. doi:10.1594/PANGAEA.72
Supplement to: **Storz, David; Schulz, Hartmut; Waniek, Joanna J; Schulz-Bull, Detlef; Kucera, Michal (2009):** Seasonal and interannual variabilit the planktic foraminiferal flux in the vicinity of the Azores Current. *Deep-Sea Research I*, **56(1)**, 107-124, doi:10.1016/j.dsr.2008.08.009

Abstract: Planktic foraminiferal (PF) flux and faunal composition from three sediment trap time series of 2002-2004 in the northeastern Atlantic show pronounced year-to-year variation despite similar sea surface temperature (SST). The averaged fauna of the in 2002/2003 is dominated by the species Globigerinita glutinata, whereas in 2003/2004 the averag fauna is dominated by Globigerinoides ruber. We show that PF species respond primarily to productivity, triggered by the seasonal dynamics of vertical stratification of the up water column. Multivariate statistical analysis reveals three distinct species groups, linked to bulk particle flux, to chlorophyll concentrations and to summer/fall oligotrophy wi SST and stratification. We speculate that the distinct nutrition strategies of strictly asymbiontic, facultatively symbiontic, and symbiontic species may play a key role in explai their abundances and temporal succession. Advection of water masses within the Azores Current and species expatriation result in a highly diverse PF assemblage. The Az Frontal Zone may have influenced the trap site in 2002, indicated by subsurface water cooling, by highest PF flux and high flux of the deep-dwelling species Globorotalia scitu Similarity analyses with core top samples from the global ocean including 746 sites from the Atlantic suggest that the trap faunas have only poor analogs in the surface sedir These differences have to be taken into account when estimating past oceanic properties from sediment PF data in the eastern subtropical North Atlantic.

Project(s): **Paleoceanography at Tübingen University** (GeoTü)

Event(s): **L1_K276** * *Latitude:* 30.0000 * *Longitude:* -22.0000 * *Elevation:* -5300.0 m * *Date/Time:* 2002-02-24T00:00:00 * *Date/Time 2:* 2004-04-01T00:00:00 * *Location:* NE Atlantic - A Front * *Device:* Trap, sediment * *Comment:* Station used since 1980

Size: 6 datasets

Download Data

Download **ZIP** file containing all datasets as tab-delimited text *(use the following character encoding:* ISO-8859-1: ISO Western (PANGAEA default) *)*

Datasets listed in this Collection

- **Storz, D; Schulz, H; Waniek, JJ et al. (2009):** (Table A a) Relative contributions of planktic foraminiferal species in sediment trap series L1/K276-22 at 2000 m water depth. doi:10.1594/PANGAEA.724294
- **Storz, D; Schulz, H; Waniek, JJ et al. (2009):** (Table A b) Flux of planktic foraminiferal species in sediment trap series L1/K276-22 at 2000 m water depth. doi:10.1594/PANGAEA.724308
- **Storz, D; Schulz, H; Waniek, JJ et al. (2009):** (Table B a) Relative contributions of planktic foraminiferal species in sediment trap series L1/K276-22 at 3000 m water depth. doi:10.1594/PANGAEA.724301
- **Storz, D; Schulz, H; Waniek, JJ et al. (2009):** (Table B b) Flux of planktic foraminiferal species in sediment trap series L1/K276-22 at 3000 m water depth. doi:10.1594/PANGAEA.724309
- **Storz, D; Schulz, H; Waniek, JJ et al. (2009):** (Table C a) Relative contributions of planktic foraminiferal species in sediment trap series L1/K276-23 at 3000 m water depth. doi:10.1594/PANGAEA.724307
- **Storz, D; Schulz, H; Waniek, JJ et al. (2009):** (Table C b) Flux of planktic foraminiferal species in sediment trap series L1/K276-23 at 3000 m water depth. doi:10.1594/PANGAEA.724310

Storz, D et al. (2009): Planktic foraminiferal flux and faunal composition of sediment trap L1_K276 in the northeastern Atlantic - Mozilla Firefox

Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

http://doi.pangaea.de/10.1594/PANGAEA.724325

Meistbesuchte Seiten Erste Schritte Aktuelle Nachrichten

Storz, D et al. (2009): Planktic forami...

Fertig

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: Storz, D et al. (2009): Planktic foraminiferal flux and faunal composition of sediment trap L1_K276 in the northeastern Atlantic. doi:10.1594/PANGAEA.724325,
Supplement to: **Storz, David; Schulz, Hartmut; Waniek, Joanna J; Schulz-Bull, Detlef; Kucera, Michal (2009):** Seasonal and interannual variability of the planktic foraminiferal flux in the vicinity of the Azores Current. *Deep-Sea Research I*, **56(1)**, 107-124, doi:10.1016/j.dsr.2008.08.009

Abstract: Planktic foraminiferal (PF) flux and faunal composition from three sediment trap time series of 2002-2004 in the northeastern Atlantic show pronounced year-to-year variations despite similar sea surface temperature (SST). The averaged fauna of the in 2002/2003 is dominated by the species Globigerinita glutinata, whereas in 2003/2004 the averaged fauna is dominated by Globigerinoides ruber. We show that PF species respond primarily to productivity, triggered by the seasonal dynamics of vertical stratification of the upper water column. Multivariate statistical analysis reveals three distinct species groups, linked to bulk particle flux, to chlorophyll concentrations and to summer/fall oligotrophy with high SST and stratification. We speculate that the distinct nutrition strategies of strictly asymbiontic, facultatively symbiontic, and symbiontic species may play a key role in explaining their abundances and temporal succession. Advection of water masses within the Azores Current and species expatriation result in a highly diverse PF assemblage. The Azores Frontal Zone may have influenced the trap site in 2002, indicated by subsurface water cooling, by highest PF flux and high flux of the deep-dwelling species Globorotalia scitula. Similarity analyses with core top samples from the global ocean including 746 sites from the Atlantic suggest that the trap faunas have only poor analogs in the surface sediments. These differences have to be taken into account when estimating past oceanic properties from sediment PF data in the eastern subtropical North Atlantic.

Project(s): **Paleoceanography at Tübingen University** (GeoTü)

Coverage: *West:* -22.0000 * *East:* -22.0000 * *South:* 30.0000 * *North:* 30.0000
Date/Time Start: 2002-02-24T00:00:00 * *Date/Time End:* 2004-03-16T00:00:00

Event(s): **L1_K276** * *Latitude:* 30.0000 * *Longitude:* -22.0000 * *Elevation:* -5300.0 m * *Date/Time:* 2002-02-24T00:00:00 * *Date/Time 2:* 2004-04-01T00:00:00 * *Location:* NE Atlantic - Azores Front * *Device:* Trap, sediment * *Comment:* Station used since 1980

Size: 6 datasets

Download Data

Download **ZIP** file containing all datasets as tab-delimited text *(use the following character encoding:* ISO-8859-1: ISO Western (PANGAEA default) *)*

Login:
Regist

Font Size:

Thumbnails | Full-Size Images

Export Citation
Add to my Quick Links
Permissions & Reprints
Cited By in Scopus (0)

hin this Article

1

ap L1/K276-22 (2000 m). (a)
butions of the 28 planktic
pecies or species varieties and

Download this File

rect

seasonal dynamics of vertical stratification of the upper water column. Multivariate statistical analysis reveals three distinct species groups, linked to bulk particle flux, to chlorophyll concentrations and to summer/fall oligotrophy with high SST and stratification. We speculate that the distinct nutrition strategies of strictly asymbiontic, facultatively symbiontic, and symbiontic species may play a key role in explaining their abundances and temporal succession. Advection of water masses within the Azores Current and species expatriation result in a highly diverse PF assemblage. The Azores Frontal Zone may have influenced the trap site in 2002, indicated by subsurface water cooling, by highest PF flux and high flux of the deep-dwelling species *Globorotalia scitula*. Similarity analyses with core top samples from the global ocean including 746 sites from the Atlantic suggest that the trap faunas have only poor analogs in the surface sediments. These differences have to be taken into account when estimating past oceanic properties from sediment PF data in the eastern subtropical North Atlantic.

Keywords: Eastern North Atlantic; Planktic foraminifers; Sediment trap; Azores Current; Particle flux; Species ecology

Article Outline

n pigments along a shipping ...
Progress in Oceanography

- An unexpected nitrate distribution in the tropical Nort...
  Deep Sea Research Part II: Topical Studies in Oceanogra...
- The physical and chemical environment and changes in co...
  Deep Sea Research Part II: Topical Studies in Oceanogra...
- High frequency and mesoscale variability in SeaWiFS chl...
  Deep Sea Research Part II: Topical Studies in Oceanogra...
- Basin-scale variability of phytoplankton bio-optical ch...
  Deep Sea Research Part I: Oceanographic Research Papers

View More Related Articles

TIB IUB Find it

Supplementary Data

View Record in Scopus

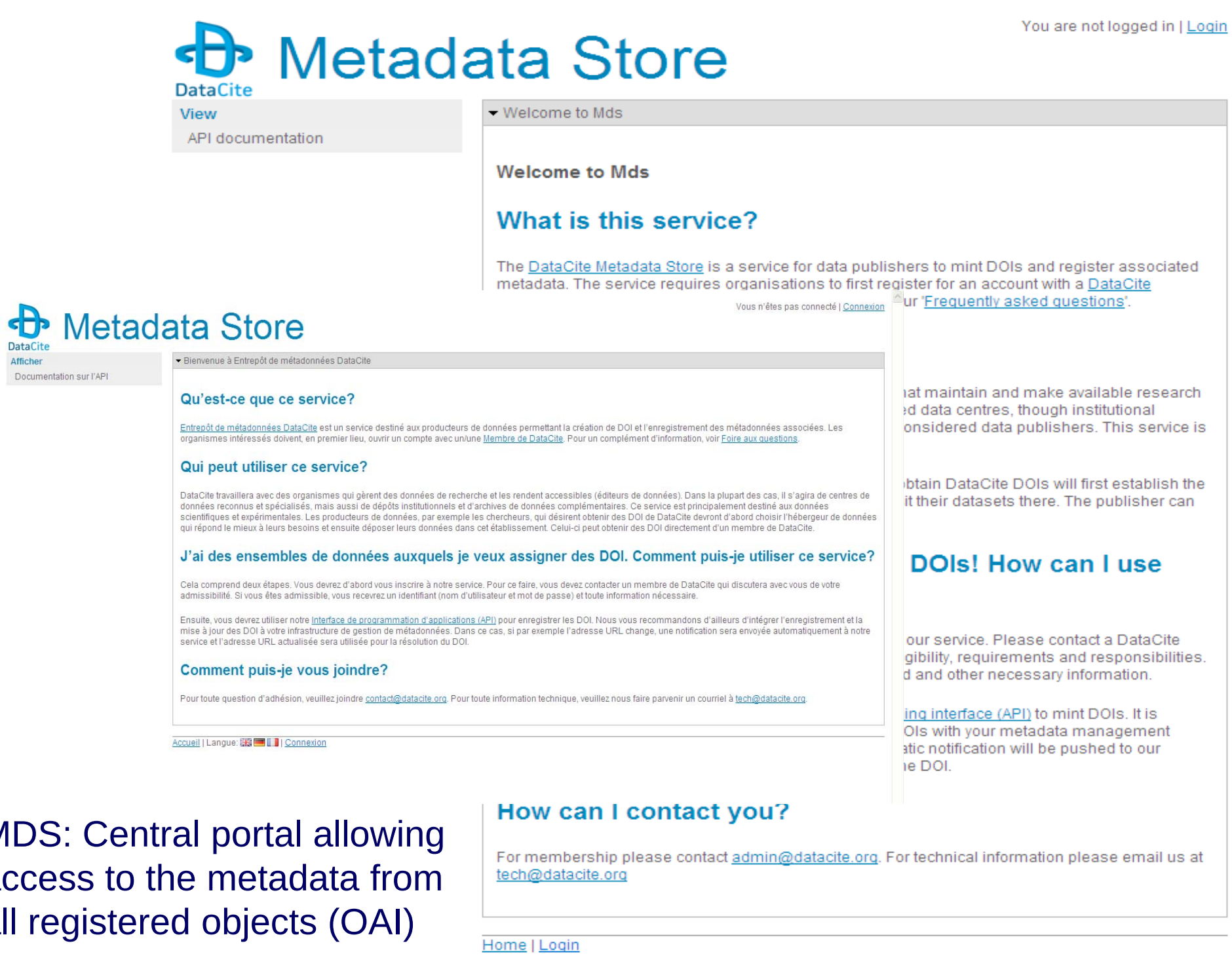
Bridging the gap

- **DataCite supports researchers by enabling them to locate, identify, and cite research datasets with confidence**
- **DataCite supports data centres by providing workflows and standards for data publication**
- **DataCite supports publishers by enabling linking from articles to the underlying data**

http://www.datacite.org
http://schema.datacite.org
https://mds.datacite.org
http://search.datacite.org
http://oai.datacite.org
http://data.datacite.org
http://stats.datacite.org

Working Groups

- Business Practices
- Criteria for Data Centers
- Identifier Syntax
- Metadata
- Services
- Special Datasets
- Technical Infrastructure



MDS: Central portal allowing access to the metadata from all registered objects (OAI)

DataCite Metadata Schema Repository

Metadata Store

For our Metadata store all versions hosted on this site are valid although only one versior different XML namespace.

Each incoming to Metadata store XML document should indicate schema location in the

Deprecation notification

DataCite Metadata Schema 2.0 will be retired on 1st Aug 2011. After this date MDS will

Available versions

- DataCite Metadata Schema 2.2 (released 2011-07-01; preferred)
- DataCite Metadata Schema 2.1 (released 2011-03-28)
- DataCite Metadata Schema 2.0 (released 2011-01-24; deprecated)

Versioning

Each available version of a schema is a subdirectory containing a set of files. Those files a pdf documentation. The files don't reference any other files outside of their version subc

Help

If you have questions please ask at Developers groups or contact tech@datacite.org.

Source code

This project is hosted on GitHub.

DataCite

DataCite - International Data Citation

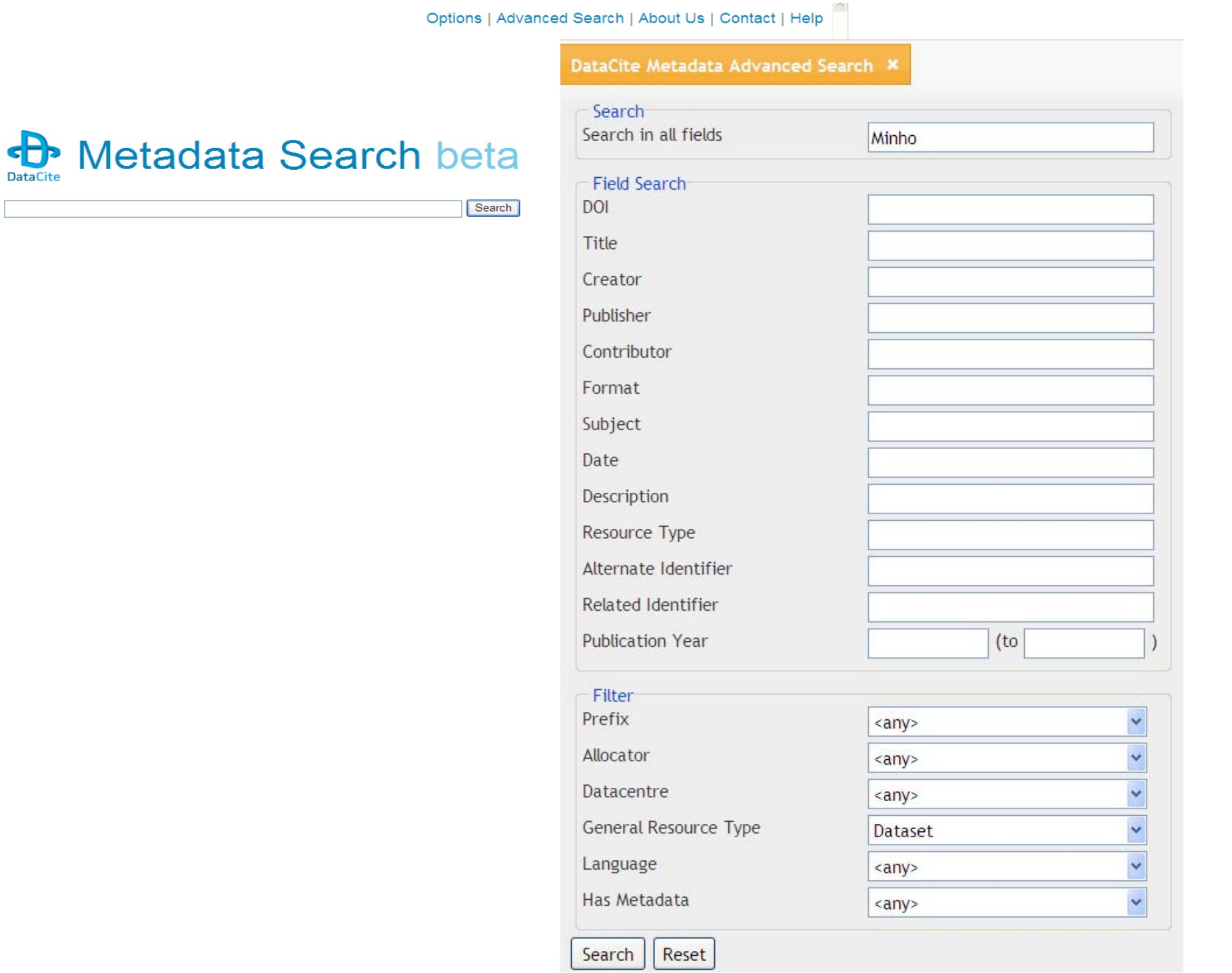
DataCite Metadata Schema for the Publication and Citation of Research Data

Version 2.2
July 2011
doi:10.5438/0005

Members of the Metadata Working Group
Joan Starr, California Digital Library (head of working group)
Jan Ashton, British Library
Jan Brase, TIB / DataCite
Paul Bracke, Purdue University
Angela Gastl, ETH Zürich
Jacqueline Gillet, Inist
Alfred Heller, DTU Library
Birthe Krog, DTU Library
Lynne McAvoy, CISTI
Karen Morgenroth, CISTI
Elizabeth Newbold, British Library
Madeleine de Smaele, TU Delft
Anja Wilde, GESIS
Scott Yeadon, ANDS
Wolfgang Zenk-Möltgen, GESIS
Frauke Ziedorn, TIB (Metadata Supervisor)

DataCite Metadata 2.2 XML Schema

```
<?xml version="1.0" encoding="UTF-8" ?>
<!--
    Revision history
        2010-08-26   Complete revision according to new common specification by the metadata work group after review. AJH, DTIC
            2010-11-17 Revised to current state of kernel review, FZ, TIB
            2011-01-17 Complete revsion after community review. FZ, TIB
            2011-03-17 Release of v2.1: added a namespace; mandatory properties got minLength; changes in the definitions of relationTypes
            IsDocumentedBy/Documents and isCompiledBy/Compiles; changes type of property "Date" from xs:date to xs:string. FZ, TIB
            2011-06-27 v2.2: namespace: kernel-2.2, additions to controlled lists "resourceType", "contributorType", "relatedIdentifierType", and "descript
-->
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns="http://datacite.org/schema/kernel-2.2" targetNamespace="http://datacite.org/schema/kernel-
  2.2" elementFormDefault="qualified" xml:lang="EN">
  <xs:include schemaLocation="include/datacite-titleType-v2.xsd" />
  <xs:include schemaLocation="include/datacite-contributorType-v2.xsd" />
  <xs:include schemaLocation="include/datacite-dateType-v2.xsd" />
  <xs:include schemaLocation="include/datacite-resourceType-v2.xsd" />
  <xs:include schemaLocation="include/datacite-relationType-v2.xsd" />
  <xs:include schemaLocation="include/datacite-relatedIdentifierType-v2.xsd" />
  <xs:include schemaLocation="include/datacite-descriptionType-v2.xsd" />
  <xs:element name="resource">
    <xs:annotation>
      <xs:documentation>Root element of a single record. This wrapper element is for XML implementation only and is not defined in the DataCite DOI standard.
        Note: This is the case for any wrapper element within this schema!</xs:documentation>
      <xs:documentation>No content in this wrapper element.</xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <!-- REQUIRED FIELDS  -->
        <xs:element name="identifier">
          <xs:annotation>
            <xs:documentation>A persistent identifier that identifies a resource.</xs:documentation>
            <xs:documentation>Currently, only DOI is allowed.</xs:documentation>
          </xs:annotation>
          <xs:complexType>
            <xs:simpleContent>
              <xs:extension base="doiType">
                <xs:attribute name="identifierType" use="required" fixed="DOI" />
              </xs:extension>
            </xs:simpleContent>
          </xs:complexType>
        </xs:element>
        <xs:element name="creators">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="creator" maxOccurs="unbounded">
                <xs:annotation>
```


Options | Advanced Search | About Us | Contact | Help
DataCite Metadata Search beta
Search
DataCite Metadata Advanced Search
Search
Search in all fields
Minho
Field Search
DOI
Title
Creator
Publisher
Contributor
Format
Subject
Date
Description
Resource Type
Alternate Identifier
Related Identifier
Publication Year
(to)
Filter
Prefix
<any>
Allocator
<any>
Datacentre
<any>
General Resource Type
Dataset
Language
<any>
Has Metadata
<any>
Search
Reset

Finding data
Overview
Catalogue search
Help on searching
About catalogue
Glossary of terms
Variables search
Browse by subject
Major studies
Major depositors
New releases
HASSET thesaurus
Other archives

About the data

Login

NEW USERS

How do I find data?
How do I register?
What data are available?

More frequently asked questions...

Print-friendly page

New Data Catalogue search

Glossary of terms in the Data Catalogue

SN 3013 -Viana do Castelo, Minho, Portugal, 1826-1931

Download/Order

Documentation **Keywords** **Publications** **Related Studies**

Title:
Viana do Castelo, Minho, Portugal, 1826-1931

Subject Categories:
Population history - History
Social history - History
Urban history - History
Population studies - Population, vital statistics and censuses
Community and urban studies - Society and culture
Elites and leadership - Social stratification and groupings

Depositor(s):
Doulton, D.C., University of Southampton

Principal Investigator(s):
Kitts, A., University of Southampton
Doulton, D.C., University of Southampton. Computing Services
Colson, F., University of Southampton. Department of History

Sponsor(s):
Economic and Social Research Council

Grant Number:
600428425022

Abstract:
Drawing on the abundant records generated by the Portuguese administrative system, the project has undertaken a reconstitution of the male urban elite of Viana do Castelo for the purpose of investigating secular demographic trends.

Main Topics:

OAI-PMH Data Provider Beta

What is this service?

This DataCite service exposes metadata stored in the DataCite Metadata Store (MDS) using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).

Who can use this service?

This service is open to everyone and is meant to be accessed by OAI-PMH compliant harvesters or any application that issues OAI-PMH requests. The service base address is `http://oai.datacite.org/oai` and the service identifier is available here.

What is OAI-PMH?

In brief, OAI-PMH provides a set of services that enables exposure and harvesting of repository metadata. The protocol is comprised of six verbs that specify the service being invoked, they are:

- **Identify** - used to retrieve information about the repository.
- **ListIdentifiers** - used to retrieve record headers from the repository.
- **ListRecords** - used to harvest full records from the repository.
- **ListSets** - used to retrieve the set structure of the repository.
- **ListMetadataFormats** - lists available metadata formats that the repository can disseminate.
- **GetRecord** - used to retrieve an individual record from the repository.

Selective harvesting can be performed by the use of accompanying parameters. Available parameters are:

- **identifier** - specifies a specific record identifier.
- **metadataPrefix** - specifies the metadata format that the records will be returned in.
- **set** - specifies the set that returned records must belong to.
- **from** - specifies that records returned must have been created/update/deleted on or after this date.
- **until** - specifies that records returned must have been created/update/deleted on or before this date.
- **resumptionToken** - a token previously provided by the server to resume a request where it last left off.

The verbs and parameters can be combined to issue requests to the service such as:

```
http://oai.datacite.org/oai?verb=Identify
http://oai.datacite.org/oai?verb=ListIdentifiers&metadataPrefix=oai_dc
http://oai.datacite.org/oai?verb=ListRecords&from=2011-06-01T00:00:00Z&metadataPrefix=oai_dc
```

For more details on the protocol, its implementation, and uses please visit the OAI-PMH web site.

Available Metadata Formats

The DataCite OAI-PMH Data Provider is able to disseminate records in the following formats:

OAI Dublin Core (oai_dc)

As a minimum requirement for OAI-PMH compliance, metadata must be made available in the OAI Dublin Core format. For more information please see the OAI-PMH web site.

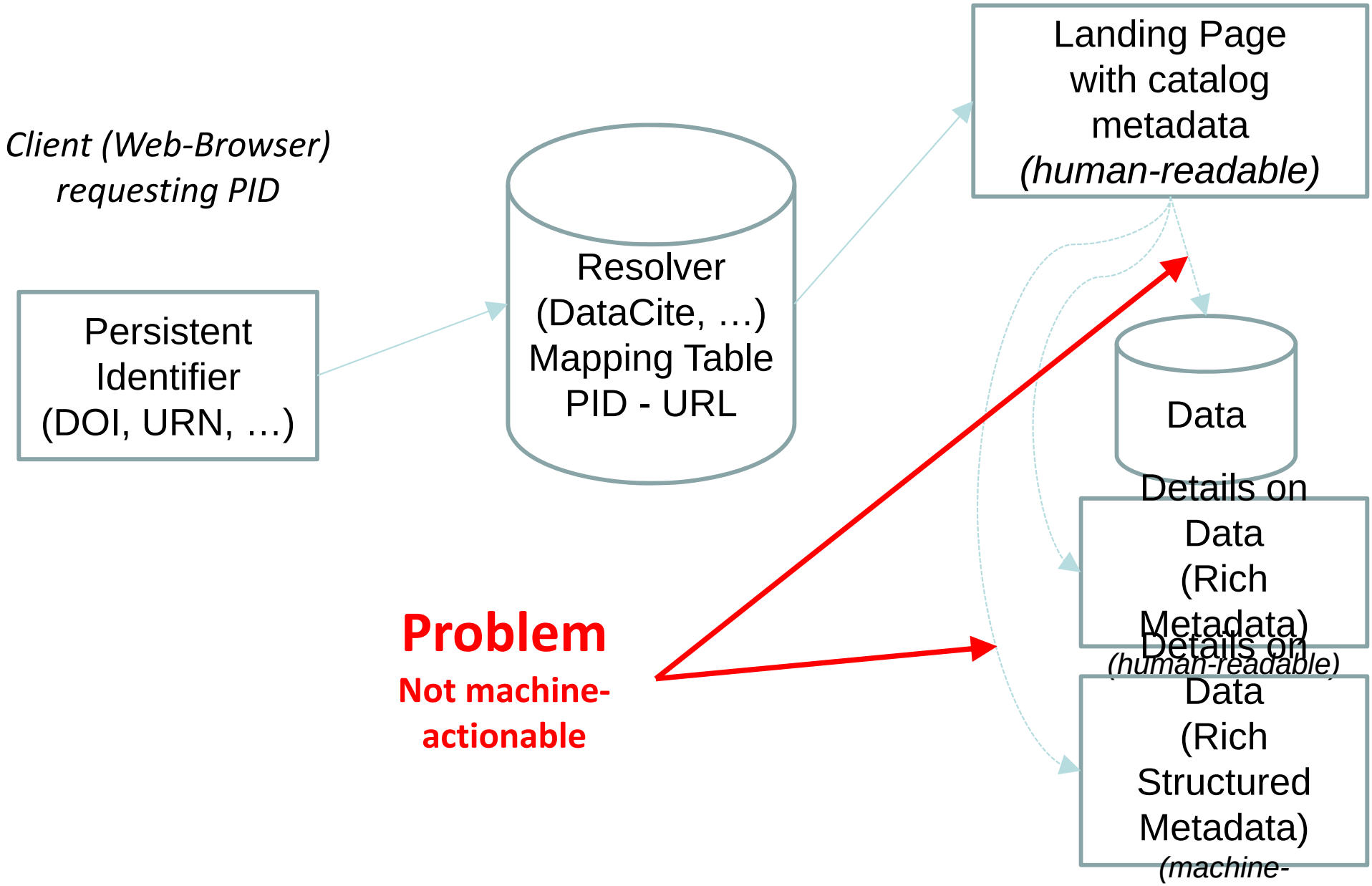
OAI DataCite (oai_datacite)

- Service for displaying DataCite metadata in different formats (BibTeX, RIS, RDF, etc.)

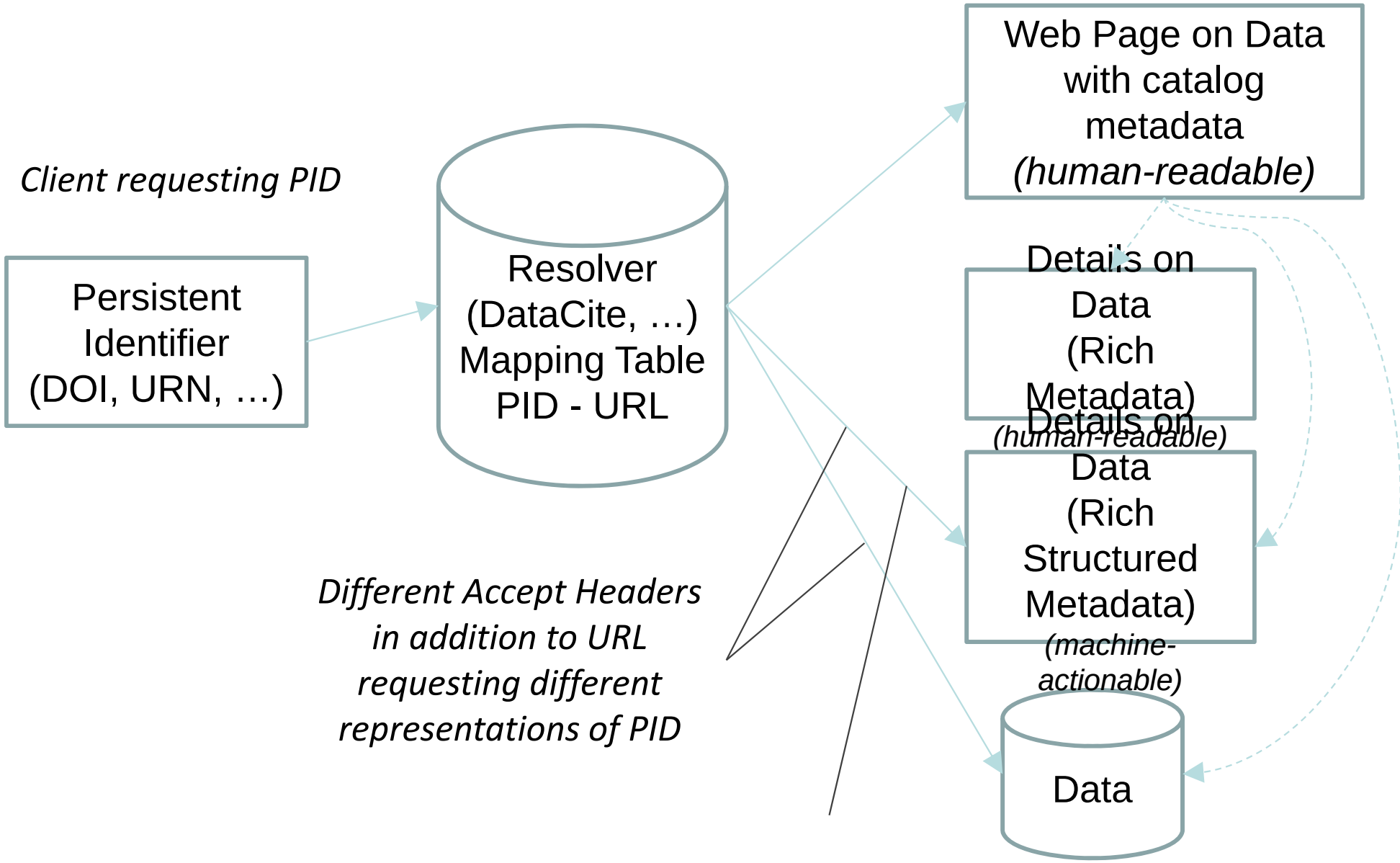
- A particular representation of the metadata can be requested via *content negotiation* or by using DOI proxy (the "http://dx.doi.org" formulation as a URL) and MIME-type

- Documentation: http://www.crosscite.org/cn/

Resolution - Current Status



Content Negotiation - Based on the Solution of CrossRef/DataCite



List of repositories for research data

Home | Members | FAQs | Services | Resources | Events | Contact us

DataCite

Helping you to find, access, and reuse data

Repositories

This list is a working document, initated via a collaboration between the British Library, BioMed Central and the Digital Curation Centre, that aims to capture the growing number of repositories for research data. It is provided for information purposes only: DataCite provides no endorsements as to the quality or suitability of the repositories listed. We encourage community participation in developing this resource. Please contact us to suggest changes or additions. A copy of the list can be downloaded from Google Docs.

Repository	Website	Subject area(s)	Funding model	Deposit restrictions	Access restrictions	Li a
Domain-specific and general data repositories (multiple formats accepted)						
3TU.Datacentrum	http://www.datacentrum	Technical sciences, climate, fluid dynamics, earth sciences	Project funded by 3TU.Federation	None, preferred formats suggested	Data is only displayed if permission has been granted by the owner who has stored it at 3TU.Datacentrum.	N li a
Access to Archival Databases	http://aad.archives.gov/	Various - social, historical, political etc	US government - part of National Archives and Records Administration (NARA)			
Archaeology Data Service	http://archaeologydatas	Archaeology	Multiple funders: AHRC, NERC, National & International project funding. ADS Charging Policy [http://archaeologydat		All data are free to access for non-commercial use [http://archaeologyda	D re Al e: to a c [h
Australian Antarctic Data Centre	http://data.aad.gov.au/	Antarctic research	Aus. Gov.		All data are freely available	
Australian Social Science Data Archive	http://www.assda.edu.a	Social, political and economic affairs	Aus. Gov.		Free access, some data may be restricted	
Biological Records Centre	http://www.brc.ac.uk/	Ecology	NERC		Free access, some data may be restricted	

Some recent related developments

- Thomson-Reuters Data Citation Index
- ORCID official launch
- ODIN European project
- CODATA/ICSTI Working Group on Data Citation
- Creation of the Research Data Alliance

THOMSON REUTERS

Search

HOME | PRODUCTS & SERVICES | NEWS & IDEAS | INVESTOR RELATIONS | ABOUT US | CAREERS

PRESS RELEASES

22 JUN 2012

THOMSON REUTERS UNVEILS DA FOR DISCOVERING GLOBAL DAT

First of Its Kind Data Citation Index Connects Repositories around the World

Philadelphia, PA, June 22, 2012 — The Intellect Thomson Reuters announced today that it will p Association Conference (ALA) the *Data Citation* resource within the *Web of KnowledgeSM* to faci attribution of data sets and data studies, and linl literature.

This new research resource from Thomson Reu discovery for scientific, social sciences and arts connecting foundational research within data re related peer-reviewed literature in journals, bool already indexed in the *Web of Knowledge*.

The Thomson Reuters *Data Citation Index*, sch makes research within the digital universe disc linked to the article detailing the outputs from the Reuters has partnered with data repositories su Consortium for Political and Social Research (IC records and cited references for digital research attribution, and ultimately the measurement of in

Thomson Reuters Speeds Research Process by Launching Data Citation Index

First-of- its-kind Data Citation Index connects researchers to data repositories around the world

THOMSON REUTERS

11/10/2012 - Philadelphia, PA, — The Intellectual Property & Science business of Thomson Reuters, the world's leading provider of intelligent information for business and professionals, announced today that it is previewing the Data Citation IndexTM at the Frankfurt Book Fair. A research resource within the Web of KnowledgeSM, the Data Citation Index facilitates the discovery, use and attribution of data sets and data studies that are foundational to peer-reviewed literature. The Data Citation Index will be available to the market by the end of October.

This new research resource from Thomson Reuters creates a single multidisciplinary source for the discovery of foundational research within data repositories around the world, placing them in the broader context of peer-reviewed literature including select journals, books, and conference proceedings already indexed in the Web of Knowledge.

Historically, it has been extremely difficult to discover, attribute and measure the research compiled within data sets. Driven by market demands, analysts at Thomson Reuters closely researched this issue in collaboration and consultation with the research community and determined that a system was needed to help researchers more easily discover and attribute data relevant to their work in a way that appropriately acknowledges intellectual credit.

"Data is not like a narrative—you don't know its value and utility until it's used. In this regard, we're doing everything we can to promote sharing of data from our archives," said John Kunze, associate director at the California Digital Library's UC Curation Center. "The Data Citation Index from Thomson Reuters is helping to turn the tide."

Thomson Reuters has collaborated with numerous data repositories worldwide to capture bibliographic records and cited references for digital research, facilitating visibility, author attribution, and ultimately the measurement of impact of this growing body of scholarship.

"We are pleased to partner with Thomson Reuters in the building of the Data Citation Index," said Mary Vardigan, assistant director of the Inter-University Consortium for Political and Social Research (ICPSR). "By enhancing the discoverability of data through the Data Citation Index, Thomson Reuters is highlighting the importance of research data in the scientific process."

l.haak@orcid.org
301-922-9062
ORCID

ORCID launches public registry for researchers

ORCID (Open Researcher and Contributor ID) is excited to announce the launch of its Registry (www.orcid.org), where researchers can distinguish themselves by creating a unique personal identifier.

"ORCID addresses a problem shared by individuals and organizations across the research community: reliably connecting research with researchers," said Laure Haak, Executive Director of the non-profit ORCID organization. "But ORCID is more than a Registry, it is a community effort to embed these identifiers in research workflows." Understanding the "who" and "what" of research has been hampered by lack of data standards, and in particular a standard for identifying individuals. Universities and other research organizations, as well as membership organizations like the American Physical Society (APS) are working to integrate ORCID identifiers into their systems. "For scholars, ORCID provides a persistent identifier that unambiguously distinguishes you as the author or creator of your published works in systems that adopt ORCID.

Through integration in workflows such as manuscript and grant submission as well as researcher profiles, ORCID promises to help scholars and institutions manage academic information and, ultimately, to provide both with more control over their own record of scholarship," said Amy Brand, Assistant Provost for Faculty Appointments and Information at Harvard University. Currently, Boston University, New York University Langone Medical Center, Cornell University, and the California Institute of Technology, and the research information system vendors Avedas, Symplectic, and Thomson Reuters are actively working on integration with the ORCID registry.

Other researcher identifier systems are currently in use serving specific populations or types of research work. ORCID does not compete with these systems, but rather provides a switchboard for crosslinking. Elsevier is providing a way for researchers to link their Scopus Author IDs to ORCID and synchronize their publication information between the two systems. Thomson Reuters' ResearcherID® will link to ORCID and allow researchers to synchronize their publication information. Several research information system providers are also planning to integrate ORCID identifiers, including KNODE, Faculty of 1000, figshare, and ImpactStory. Through its affiliate ORCID EU, ORCID is working with DataCite to link ORCID identifiers with research datasets.

"ORCID is unique in its ability to reach across disciplines, research sectors, and national boundaries, and has great potential for solving the name ambiguity problem in scholarly communications," explained Howard Ratner, Chair of ORCID and Chief Technology Officer of Nature Publishing Group. "Publishers have used several different naming schemes for authors and contributors. Commonly occurring names and name changes during a researcher's career has led to difficulties linking authors and contributors with their works." Nature Publishing Group is among a group of Launch Partners, along with Hindawi Publishing Corporation, Aries Systems, Thomson Reuters, and the American Physical Society (APS), that are integrating ORCID identifiers into the manuscript submission process.

In addition to research organizations, higher education institutions and publishers, ORCID is working with research funders including The Wellcome Trust and the National Institutes of Health to explore how ORCID identifiers might be linked into the grant application process. Liz Allen, Evaluation lead at The Wellcome Trust, notes, "An ongoing challenge for funders is tracking the outputs and outcomes of research we have funded. Over time, the Wellcome Trust plans to integrate ORCID identifiers into our online grant application and reporting systems so we can easily link researchers to works they create and reduce reporting burdens for researchers while helping us to better understand the impact of our funding."

###

ORCID, was established in 2010 as non-profit organization serving the research community. For more information on ORCID, please see http://orcid.org.

ORCID and DataCite Interoperability Network

"Our vision is a scientific e-infrastructure that supports seamless access, use, re-use, and trust of data. In a sense, the physical and technical infrastructure becomes invisible and the data themselves become the infrastructure – a valuable asset, on which science, technology, the economy and society can advance"

(Riding the wave. the High Level Expert Group on Scientific Data, 2010)

'Data as infrastructure' is a critical concept for a fully-integrated European Research Area (ERA) to drive innovation forward as envisaged by the Digital Agenda for Europe. The lack of data availability hinders this vision. In academic publishing, peer review and citation have long been recognised as mechanisms for endorsing the trustworthiness of research outputs and incentivizing researchers to contribute. Trustworthy research data will only be widely available if the same principles are applied. Key, participative, initiatives have emerged to address this challenge.

The DataCite consortium has assigned over 1m DOI names in the last few years to make research data citable, true to emergence of the '4th paradigm', Jim Gray's vision of "data-intensive scientific discovery".

ORCID offers the opportunity to identify individual authors across systems and infrastructures, including scholarly works that can have up to thousands of authors (as in the case of the LHC project). Researchers often change their affiliation, and collaborate across national disciplinary boundaries.

ODIN aims to build on the success of DataCite and ORCID by designing an 'awareness layer' for persistent author and object identifiers, thereby reducing technical, cultural and logistical barriers to the accessibility, attribution and trust of data. Identifier awareness will make it possible to stabilise:

- References to a data object
- Tracking of use and re-use
- Links between a data object, subsets, articles, rights statements and every person involved in its life-cycle (creator, editor, reviewer, aggregator, etc.).

Given the importance of these functions as we approach HORIZON2020, we aim to prove the feasibility of author, data and rights identification, promote trust building towards open scientific data e-Infrastructures and lay the foundation necessary to promote future interoperability (technical, semantic, reference architecture, etc.) in the scientific data domain in Europe and globally.

« ODIN will build on the ORCID and DataCite initiatives to uniquely identify scientists and data sets and connect this information across multiple services and infrastructures for scholarly communication.

It will address some of the critical open questions in the area: Referencing a data object; Tracking of use and re-use; Links between a data object, subsets, articles, rights statements and every person involved in its life-cycle. »

Data Citation Standards and Practices Task Group

Approved by the CODATA 27th General Assembly in Cape Town 2010

The need for robust data citation capabilities

As the growth of electronic publishing of literature has created new challenges, such as the need for mechanisms for citing online references in ways that can be relied upon for many years into the future, the growth in online datasets (as distinguished from literature) presents related, yet additional challenges. Data citation standards and good practices can also form the basis for increased incentives, recognition, and rewards for scientific data activities that in many cases are currently lacking in all fields of research. The rapidly-expanding universe of online digital data holds the promise of allowing peer-examination and review of conclusions or analysis based on experimental or observational data, as well as the ability for subsequent users to make new and unforeseen uses and analyses of the same data – either in isolation, or in combination with other datasets.

This promise, however, depends upon the ability to reliably identify, locate, access, interpret and verify the version, integrity, and provenance of the digital datasets. The problem of citing online data is complicated by the lack established practices for referring to portions or subsets of data. Unlike in the realm of literature where a printed edition may be the version of record for a document, typically there is no such hard-copy of a database. Even if it were feasible for there to be such a hard-copy, scientists lack the necessary constructs and conventions for referring to portions of a database, analogous to the volume and page numbers, or titles, chapters, and sections, that we use commonly in citing to text that is published in books or serial publications.

As funding sources for scientific research have begun to require data management plans as part of their selection and approval processes, it is important that the necessary standards, incentives, and conventions to support data citation, preservation, and accessibility be put into place. There are, in fact, a number of initiatives in different organizations, countries, and disciplines already underway. One important group is DataCite. Others remain ad hoc and uncoordinated. The Task Group, being organized jointly by several CODATA committees and ICSTI, together with representatives from several other organizations, would examine a number of key issues related to data citation, help coordinate activities in this area, and promote common practices and standards in the scientific community.

Issues Requiring Attention

There are many issues that need to be addressed in establishing standards and good practices in the data citation arena. Below is a preliminary and partial annotated list that the Task Group would consider, prioritize, and address as appropriate.

A. Technical

1. Interoperability and Facilitation of Re-use. There is already considerable diversity in database formats, such as various flat-file, hierarchical, relational, object-oriented, and XML-based databases. There is every reason to expect that new modalities and formats for storing and manipulating digital data will continue to emerge.
2. Citation Formats. What data citation conventions have been developed already? How are they similar and how do they differ? Can they be standardized?
3. Metadata. How do metadata conventions or standards affect citation formats?
4. Database Versioning. Datasets are more dynamic than documents, and this creates additional challenges for citation practice. When should the dataset as a whole be cited? How can a specific, time-fixed version be cited? What changes to the data constitute a new contribution or added value? How should this be acknowledged? How are database versions controlled and labelled?

http://www.codata.org/taskgroups/TGdatacitation/index.html

Thank you


Knowledge
for Life
ZB MED


gesis
Leibniz Institute
for the Social Sciences


ZBW
Leibniz-Informationszentrum
Wirtschaft
Leibniz Information Centre
for Economics


TIB
TECHNISCHE
INFORMATIONSBIBLIOTHEK


สำนักงาน
คณะกรรมการวิจัยแห่งชาติ
วช.
NRCT
National Research
Council of Thailand

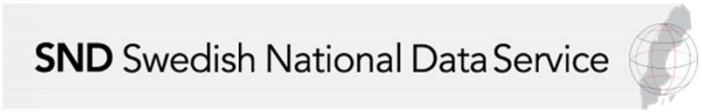

NRC·CNRC
CISTI
ICIST


DTU


inist


TU Delft


KiSTi
www.kisti.re.kr


SND Swedish National Data Service


华大基因
BGI


ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich


BRITISH
LIBRARY


DCC


University of California
CDL
California Digital Library


osti.gov


PURDUE
UNIVERSITY
LIBRARIES


ands
AUSTRALIAN NATIONAL DATA SERVICE


ICPSR
A PARTNER IN
SOCIAL SCIENCE
RESEARCH
INTER-UNIVERSITY
CONSORTIUM FOR
POLITICAL AND
SOCIAL RESEARCH


Microsoft
Research


CRUI
Conferenza dei Rettori
delle Università Italiane