



# Les données dans l'écosystème des publications scientifiques

Odile Hologne, Directrice déléguée à l'IST

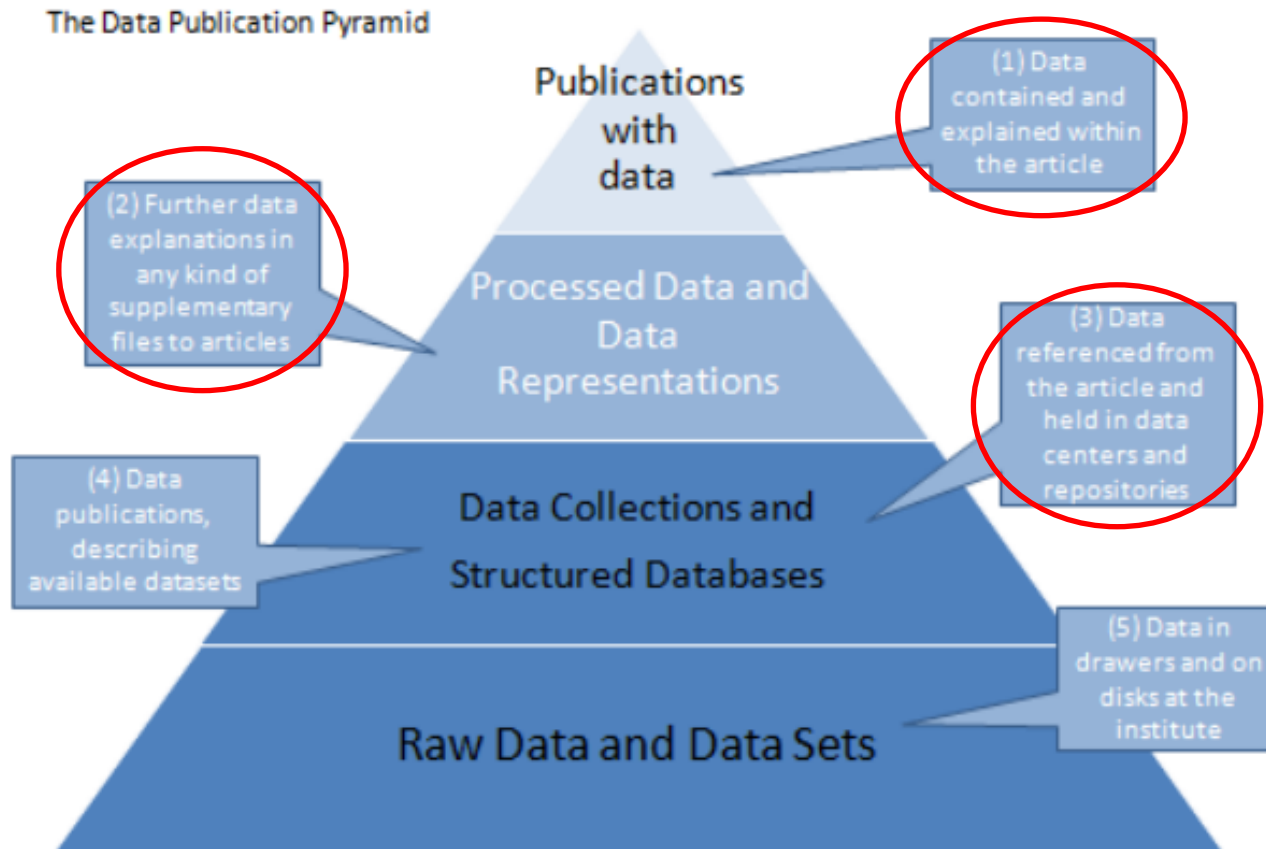


@Holo\_08

# Plan



- ❖ Préambule : données, H2020, acteurs
- ❖ Revues scientifiques et données
- ❖ Lien articles-données : entrepôts
- ❖ Citation des articles, citation des données
- ❖ Stratégie des grands éditeurs
- ❖ Conclusion



Graph 1: The Data Publication Pyramid, developed on the basis of the Jim Gray pyramid, to express the different manifestation forms that research data can have in the publication process.



# OA to publications mandate in H2020

**Each beneficiary must ensure OA to all peer-reviewed scientific publications relating to its results:**

- **Deposit** a machine-readable copy of the published version or final peer-reviewed manuscript accepted for publication in a repository of the researchers choice (possibly OpenAIRE compliant)
- **Ensure OA** on publication or at the latest within 6 months (12 for SSH)
- **Aim to deposit** at the same time **the research data needed to validate the results ("underlying data")**
- Ensure **OA to the bibliographic metadata** that identify the deposited publication, via the repository

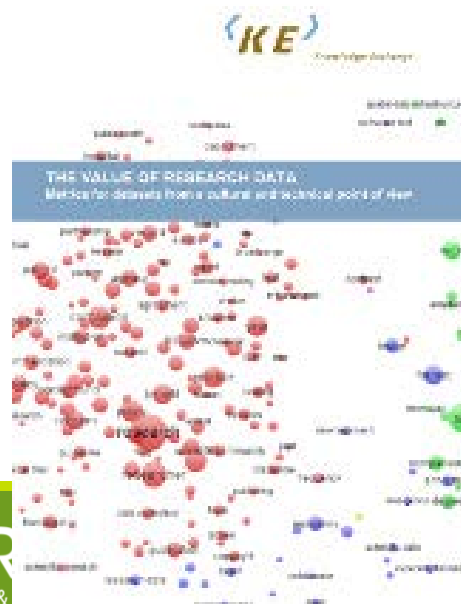
Celina Ramjouw (Head of Sector "Open Access to scientific Publications and Data", EC DG CNECT)

Agenda

# Data sharing : different stakeholders

Knowledge exchange Report : the value of research data

<http://www.knowledge-exchange.info/datametrics>



## Recommendations for the different stakeholders

Stakeholders	Recommendations
Funders	<ul style="list-style-type: none"> <li>* Demand and reward data sharing activities</li> <li>* Consider data metrics in assessments</li> <li>* Inform policy about the importance and benefits of data sharing</li> <li>* Promote open access of data</li> </ul>
Research infrastructures	<ul style="list-style-type: none"> <li>* Promote policies of data sharing</li> <li>* Promote arguments and incentives in favour of data sharing</li> <li>* Provide options and alternatives to the different types of data sharing activities</li> <li>* Professionalize staff and standardize data sharing activities (collection, curation, dissemination)</li> </ul>
Scientists	<ul style="list-style-type: none"> <li>* Include data sharing as good scientific and scholarly practice</li> <li>* <u>Promote data citation as the formal way of acknowledging data sharing</u></li> <li>* Perform more research on benefits and possibilities of data sharing</li> <li>* Define codes of conducts for disciplines considering appropriate regulations, i.e. embargo periods, anonymisation etc.</li> </ul>
Data centres	<ul style="list-style-type: none"> <li>* Inform the scientific community about data activities and services</li> <li>* Contribute to reduce the dispersion of data repositories</li> <li>* Develop robust solutions for the preservation and standardisation of the data storage and citations</li> <li>* Develop tools for tracking the users of the repositories</li> </ul>
Publishers	<ul style="list-style-type: none"> <li>* <u>Promote data sharing in their publications and journals</u></li> <li>* Inform authors about other data sharing stakeholders (e.g. repositories, data centres)</li> <li>* <u>Support open access to data</u></li> </ul>
Libraries	<ul style="list-style-type: none"> <li>* Promote data publications and data citations</li> <li>* Coach scholars and research managers in their data publication and citation activities</li> <li>* Inform authors about other data sharing stakeholders (e.g. funders, repositories, data centres)</li> <li>* Develop tools to find data repositories</li> <li>* Develop and test appropriate metrics</li> </ul>
Publication databases	<ul style="list-style-type: none"> <li>* Collect and measure data publications and data citations</li> <li>* Facilitate the analysis and metrics of data publications and data citations</li> </ul>



# Revue scientifique et données

Nouvelles revues, évolution des revues traditionnelles

# Données et publications

## ❖ Des nouvelles revues :

(GIGA)<sup>n</sup>  
SCIENCE

BGI Shenzhen and  
BioMed Central

GigaScience aims to revolutionize data dissemination, organization, understanding, and use. An online open-access open-data journal, **we publish 'big-data' studies from the entire spectrum of life and biomedical sciences.** To achieve our goals, the journal has a novel publication format: one that **links standard manuscript publication with an extensive database that hosts all associated data and provides data analysis tools and cloud-computing resources.**

**Not just 'omic' type data** ... imaging, neuroscience, ecology, cohort data, systems biology and other new types of large-scale sharable data.

## ❖ Note aux auteurs – revues classiques

**Science**

**Data and materials availability** All data necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of *Science*. ]...[  
*Science* supports the efforts of databases that aggregate published data for the use of the scientific community. Therefore, appropriate data sets (including microarray data, protein or DNA sequences, atomic coordinates or electron microscopy maps for macromolecular structures, and climate data) must be deposited in an approved database, and an accession number or a specific access address must be included in the published paper. We encourage compliance with MIBBI guidelines (Minimum Information for Biological and Biomedical Investigations).

# Data journals, data articles

A growing list :

<http://proj.badc.rl.ac.uk/preparde/blog/DataJournalsList>

A data paper describes a dataset, giving details of its collection, processing, software, file formats etc, without the requirement of novel analyses or ground breaking conclusions. It allows the reader to understand the when, how and why data was collected and what the data-product is



## Resolving the publishing bottleneck for biodiversity

*Science is a combination of gathering facts and making theories; neither can progress on its own. In the history of science, the laborious accumulation of facts is the dominant mode, not a novelty.*

Peter Norvig



### Making "small" data big

- No lower/upper limit of manuscript size
- Publish all kinds of biodiversity related data
- Reduced page charges affordable by all



### More than just data journal!

- Integrated text and data publishing
- Completely online revisions and editing
- Community ownership of data

Free of charge in launch phase



### Community peer-review

- 7 weeks from submission to decision
- 3 days from acceptance to publication
- Public peer-review on author's choice



# Data article



Biodiversity Data Journal 2: e4123 (30 Oct 2014)  
doi: 10.3897/BDJ.2.e4123



## A database of life-history traits of European amphibians

▼ [Audrey Trochet](#), [Sylvain Moulherat](#), [Olivier Calvez](#), [Virginie M Stevens](#), [Jean Clobert](#), [Dirk S Schmeller](#)

### Abstract

In the current context of climate change and landscape fragmentation, efficient conservation strategies require the explicit consideration of life history traits. This is particularly true for amphibians, which are highly threatened worldwide, composed by more than 7400 species, which constitute one of the most species-rich vertebrate groups. The collection of information on life history traits is difficult due to the ecology of species and remoteness of their habitats. It is therefore not surprising that our knowledge is limited, and missing information on certain life history traits are common for in this species group. We compiled data on amphibian life history traits from literature in an extensive database with morphological and behavioral traits, habitat preferences and movement abilities for 86 European amphibian species (50 Anuran and 36 Urodela species). When it were available, we reported data for males, females, juveniles and tadpoles. Our database may serve as an important starting point for further analyses regarding amphibian conservation.

### Keywords

Amphibians, life history traits, Europe

Contents

Article info

Citation

Metrics

Figures

Tables

Taxa

Data

References

[Article title](#)

[Abstract](#)

[Introduction](#)

[General description](#)

[Geographic coverage](#)

[Taxonomic coverage](#)

[Traits coverage](#)

— [Morphological traits](#)

— [Life history traits](#)

— [Movement](#)

— [Habitat preferences and distribution](#)

— [Threats](#)

[Usage rights](#)

[Data resources](#)

[Additional information](#)

[Acknowledgements](#)

[Author contributions](#)

[References](#)

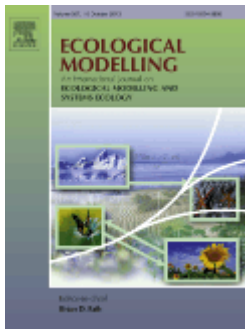
# Journaux classiques : note aux auteurs

## Incitation au dépôt des données

**Science**

**Data and materials availability** All data necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of *Science*. ]...[

*Science* supports the efforts of databases that aggregate published data for the use of the scientific community. Therefore, appropriate data sets (including microarray data, protein or DNA sequences, atomic coordinates or electron microscopy maps for macromolecular structures, and climate data) must be deposited in an approved database, and an accession number or a specific access address must be included in the published paper. We encourage compliance with MIBBI guidelines (Minimum Information for Biological and Biomedical Investigations).



### 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](https://doi.org/10.1016/0016-7037(95)00105-9). Please use PANGAEA's web interface to submit your data (<http://www.pangaea.de/submit/>).

### Data deposition

Large-scale datasets, sequences, atomic coordinates and computational models should be deposited in one of the relevant public databases prior to submission (provided private access is available at the database) and authors should include accession codes in the Materials & Methods section. The suggested wording for referring to accession identifiers in a manuscript is the following: "The [protein interaction | microarray | mass spectrometry ] data from this publication have been submitted to the [name of the database] database [URL] and assigned the identifier [accession | permalink | hashtag ]." If necessary, please include in the manuscript the relevant information (username and password) for confidential access by peer-reviewers.



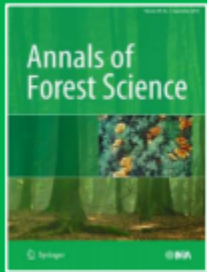
## ❖ Survey of journals author guidelines

Results of Journal Survey	
Total no. of Journals surveyed	371
Total no. of Journals with data sharing policies	162
Total no. of Journals that make sharing a requirement of publication	31
Total no. of Journals that enforce the policies	27
Total no. of Journals that state consequences for non compliance	7

The JoRD Project was a feasibility study on the possible shape of a central service on journal research data policies, funded by the UK JISC under its Managing Data Research Programme

<http://jordproject.wordpress.com/category/data-analysis/>

# Journaux classiques : data papers



## Annals of Forest Science

[http://www.springer.com/cda/content/document/cda\\_downloaddocument/IFA-AFS-June2014.pdf?SGWID=0-0-45-1073837-p174076259](http://www.springer.com/cda/content/document/cda_downloaddocument/IFA-AFS-June2014.pdf?SGWID=0-0-45-1073837-p174076259)

• *Data papers*: Data bases play an increasing role in forest and ecological sciences. While some large databases are currently available for data like forest production and inventories, plant and animal communities, carbon and nutrient cycles, there are many fields of research that produce data bases that are not publicly available and remain unknown. Such databases remain frequently under-valued and analysed while they contain information that may be of use for further analysis, novel questions and meta-analyses. AFS therefore decided to create a journal section devoted to data bases. A data paper should provide a clear and exhaustive description of the database, of its content and its potential use and all the metadata required to access the data base. Additional information will be made available about how to implement the preparation of such papers. Researchers using the data-base will be able to cite the data paper and therefore acknowledge the origin of the data; they may also invite the data-base providers as co-authors to the new publication.

A data paper should be arranged as follows:

(i) title without author's names and affiliations; (ii) introduction presenting the main features and the potential use of the data base; (iii) brief description of the structure of the data-base and of the metadata (the detailed meta-data will be provided in a specific file); (iv) electronic access to the data base and conditions for the use of the data; (v) conclusion (vi) references, (vii) tables if any; (viii) captions of figures and (ix) figures if any.

# Intérêt des Data papers

- ❖ Un article : compatible avec le dispositif d'évaluation actuel
- ❖ Documente les données
- ❖ Contribue à leur « découvrabilité »
- ❖ Favorise la citation
- ❖ Donne une place aux technologies de l'info dans les sciences



# Entrepôt de données – lien article-données

Où déposer ses données ?

# Dépôt des données : 2 approches

## ❖ Supplementary materials

- ✓ certainement pas la meilleure méthode

## ❖ Dépôt dans un entrepôt (métadonnées + data sets)

- ✓ disciplinaire
- ✓ généraliste
- ✓ institutionnel



# Entrepôts de données



**Dryad is an international repository of data** underlying peer-reviewed articles in the basic and applied biosciences. Dryad enables scientists to validate published findings, explore new analysis methodologies, repurpose data for research questions unanticipated by the original authors, and perform synthetic studies. Dryad is governed by a consortium of journals that collaboratively promote data archiving and ensure the sustainability of the repository.



The information system PANGAEA is operated as an Open Access library aimed at archiving, publishing and distributing **georeferenced data from earth system research**. The system guarantees long-term availability of its content through a commitment of the operating institutions.

Most of the data are freely available and can be used under the terms of the license mentioned on the data set description. A few password protected data sets are under moratorium from ongoing projects. The description of each data set is always visible and includes the principle investigator (PI) who may be asked for access. Each dataset can be identified, shared, published and cited by using a Digital Object Identifier (DOI). Data are archived as supplements to publications or as citable data collections. Citations are available through the portal of the German National Library of Science and Technology



ESA Data Registry is a publicly accessible registry describing scientific data sets on **ecology and the environment**. The data sets registered here are associated with articles published in the journals of the Ecological Society of America. They are registered here in order to facilitate communication and data sharing by scientists.





### Characterization of the complete chloroplast genome of *Hevea brasiliensis* reveals genome rearrangement, RNA editing sites and phylogenetic relationships

Sithichoke Tangphatsomruang, Pichahpuk Uthapaisanwong, Duangjai Sangsrakru, Juntima Chanprasert, Thippawan Yoocha, Nukoon Jomchai, Somvong Traogoonrung

Show more

doi:10.1016/j.chemgeo.2011.01.002

Get rights and content

#### Abstract

Rubber tree (*Hevea brasiliensis*) is an economical plant and widely grown for natural rubber production. However, genomic research of rubber tree has lagged behind other species in the Euphorbiaceae family. We report the complete chloroplast genome sequence of rubber tree as being 161,191 bp in length including a pair of inverted repeats of 26,810 bp separated by a small single copy region of 18,362 bp and a large single copy region of 89,209 bp. The chloroplast genome contains 112 unique genes, 16 of which are duplicated in the inverted repeat. Of the 112 unique genes, 78 are predicted protein-coding genes, 4 are ribosomal RNA genes and 30 are tRNA genes. Relative to other plant chloroplast genomes, we observed a unique rearrangement in the rubber tree chloroplast genome: a 30-kb inversion between the *tmE(UUC)-tmS(GCU)* and the *tmT(GGU)-tmR(UCU)*. A comparison between the rubber tree chloroplast genes and cDNA sequences revealed 51 RNA editing sites in which most (48 sites) were located in 26 protein coding genes and the other 3 sites were in introns. Phylogenetic analysis based on chloroplast genes demonstrated a close relationship between *Hevea* and *Manihot* in Euphorbiaceae and provided a strong support for a monophyletic group of the eusoid I.

### the Walvis Ridge basalts

Get rights and content

... are many and various and some require an...
... ment are many and various and some require an...
... diment or an enriched melt from the subcontinental...
... ality' for the Enriched Mantle I (EMI) component. We...
... te 528 on the Walvis Ridge with the aim to constrain...
... measured for the three sites overlap with the values...
... is less radiogenic Sr and more radiogenic Pb and Nd...
... ent and radiogenic isotope (Hf, Nd, Pb and Sr)...
... o produce the simplest possible model that satisfies...
... elevated <sup>207</sup>Pb/<sup>204</sup>Pb with respect to <sup>206</sup>Pb/<sup>204</sup>Pb...
... sed origins had modeled it as such. The data is...
... re addition of a melt to a mantle with bulk Earth-like...
... melt. The timing of these two events is such that the...
... d the subsequent melt removal before 3.5 Ga. This...
... EMI component.

... ent mixing trends: one formed by the less enriched...
... Site 525A basalts. The two trends have the EMI...
... Site 527–Site 528 basalts is FOZO-like and can be...
... saltic oceanic crust plus sediment). This recycled...
... nd magnitude of the chemical fractionation resulting...
... ydration experiments on basalts and sediment.

... basalts have flatter REE patterns and show less...
... EE elements. Using the isotopic compositions as...
... o model the trace element patterns of the basalts as

#### Recommended articles

Geochemistry of sulfides in Hawaiian garnet pyrox...

2010, Chemical Geology more

Growth of complex sheeted zones during recyclin...

2008, Journal of Volcanology and Geothermal Research more

Tracing fluid–rock reaction and hydrothermal circu...

2010, Chemical Geology more

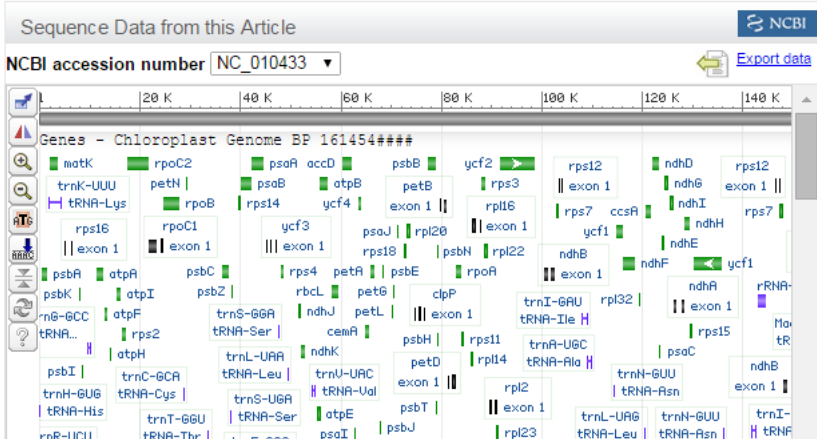
View more articles »

#### Citing articles (19)

#### Related book content

PANGAEA® – Related Data
Geochemistry of Walvis Ridge basalts
Map Satellite
Google
Map Data 1000 km Terms of Use

#### Data for this Article





# (GIGA)<sup>n</sup> SCIENCE

Journal, data-platform and database for large-scale data



in conjunction with



## Open-Paper



[DOI:10.1186/2047-217X-1-18](https://doi.org/10.1186/2047-217X-1-18)

Highly accessed

>11000 accesses

Linked to  
DOI

(GIGA)<sup>n</sup><sub>DB</sub>  
Data sets

**Open-Data**  
[DOI:10.5524/100038](https://doi.org/10.5524/100038)

78GB CC0 data

DOI

Linked to

(GIGA)<sup>n</sup> Galaxy  
by CBIIT  
Analyses

**Open-Pipelines**  
**Open-Workflows**

[DOI:10.5524/100044](https://doi.org/10.5524/100044)



# Citation articles et citations des données

Corrélation, mesure d'audience

## Respectful re-use

*Nature Genetics* **44**, 1073 (2012) | doi:10.1038/ng.2433

Published online 26 September 2012

**The impact of the papers we publish depends increasingly on the data they describe. In insisting on data access for referees and readers, we prioritize scientific integrity above all and place the interests of research participants before impact.**

The Nature research journals insist that materials, methods and data be made available and that authors detail any conditions for access where these exist. In our shared guide to authors (<http://www.nature.com/authors/policies/availability.html>), we state:

*The preferred way to share large data sets is via public repositories. Some of these repositories offer authors the option to host data associated with a manuscript confidentially, and provide anonymous*



print



email



pdf



download citation



order reprints



rights and permissions

# Lien avec les citations des articles

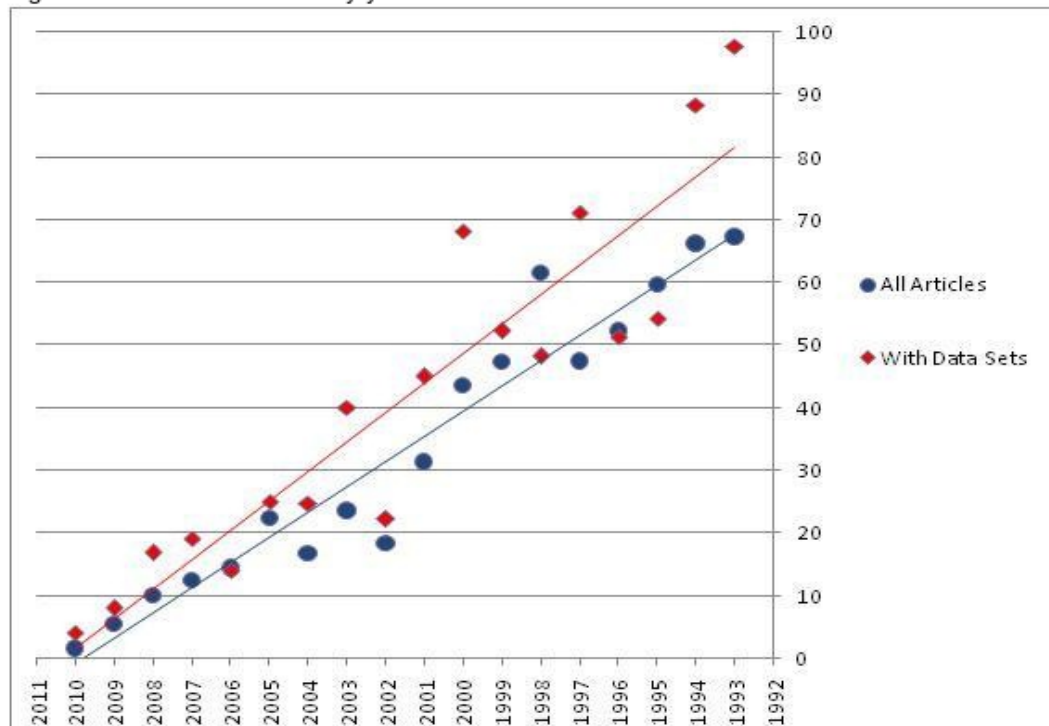
## Data Sharing Effect on Article Citation Rate in Paleoceanography

posted Nov 27, 2011 4:11 AM by Michael Diepenbroek [ updated Nov 27, 2011 11:21 PM ]

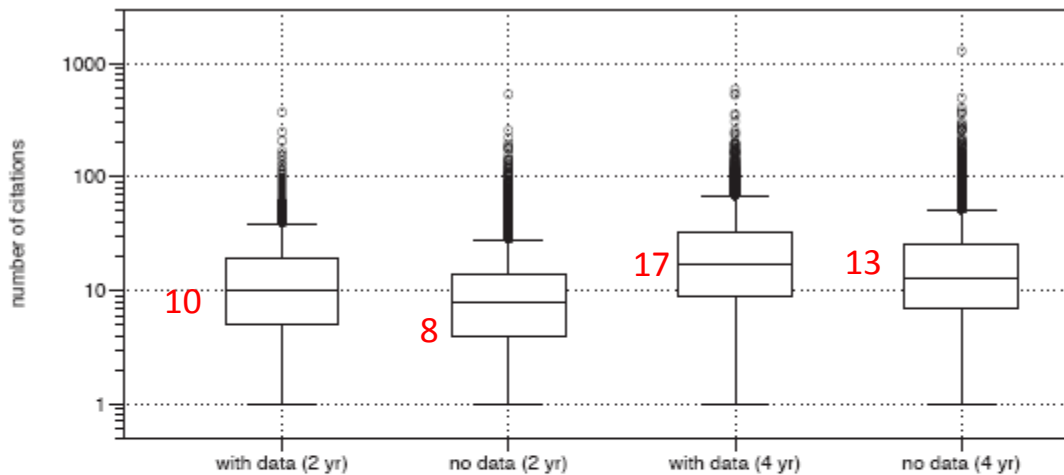
Jon Sears (AGU) - Abstract for the AGU 2011:

The validation of scientific results requires reproducible methods and data. Often, however, data sets supporting PANGAEA® data library measurably increases the citation rate of articles published between 1993 and 2010 in available supporting data sets received 19.94% (8,056) of the aggregate citations (40,409). Publicly available citations sampled over the 18-year study period (1,331), and the increase is fairly consistent over time (14 of 18 years).

Figure: Plot of mean citations by year to articles with data sets vs. citations to all articles in Paleoceanography between



# Link to data and effect on citation



<http://arxiv.org/pdf/1111.3618v1.pdf>

Linking to Data - Effect on Citation Rates in Astronomy

Figure 1. Distribution of citations of articles published in *The Astrophysical Journal* (including *Letters* and *Supplement*), *The Astronomical Journal*, *The Monthly Notices of the R.A.S.* and *Astronomy & Astrophysics* including *Supplement*), during the period 1995 through 2000. The extent of the box corresponds with the interquartile range of the citations and whiskers extend to 1.5 times the horizontal lines within the boxes correspond with the media boxes correspond respectively with the citation distribution and without data links 2 years after publication, and 4 years medians are respectively at 10, 8, 17 and 13 citations.

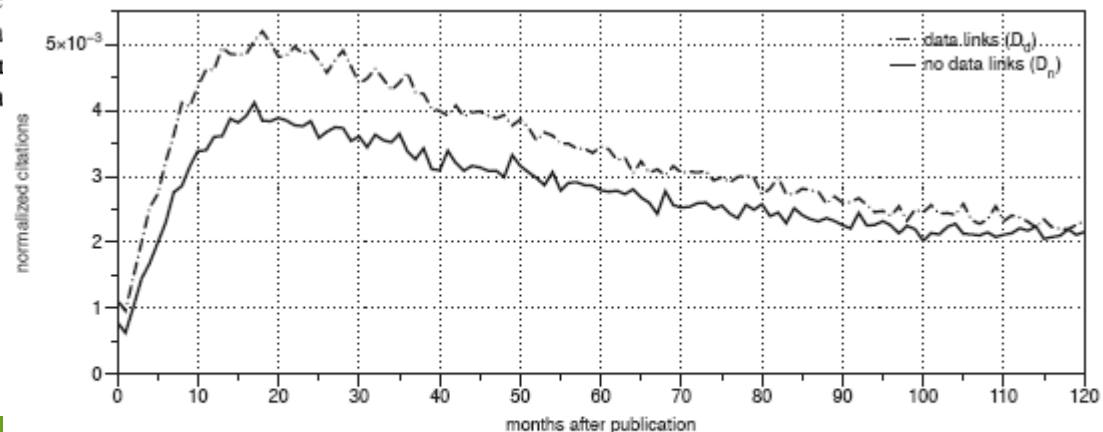


Figure 2. The normalized number of citations for data sets  $D_d$  and  $D_n$ . The citations have been normalized by the total number of citations.



# Building a Culture of Data Citation




# THE DATA CITATION INDEX<sup>SM</sup>

## DEFINITIONS:

**Data repository:** a database or collection comprising data studies, data sets and/or microcitations which stores and provides access to the raw data. Constituent data studies, and sometimes individual data sets, are marked up with metadata providing a context for the available raw data.

**Data study:** description of studies or experiments held in repositories with the associated data which have been used in the data study. (Includes serial or longitudinal studies over time). Data studies can be a citable object in the literature and may have cited references attached in their metadata, together with information on such aspects as the principal investigators, funding information, subject terms, geographic coverage etc. The level of metadata provided varies between repositories.

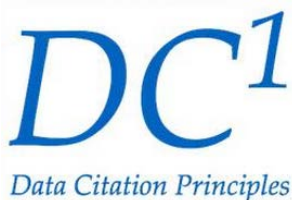
**Data set:** a single or coherent set of data or a data file provided by the repository, as part of a collection, data study or experiment. Data sets may present in a number of file formats and media types: they may be number based files such as spreadsheets, images, video, audio, databases etc. Data sets can be a citable object in the literature and may have cited references attached in their metadata, but more commonly they inherit the metadata of the overall study in which they are used.

- 1. Title: **ESTerases and alpha/beta Hydrolase Enzymes and Relatives.**  
Editor(s): Hotelier, Thierry; Renault, Ludovic; Cousin, Xavier; et al.  
Source: ESTerases and alpha/beta Hydrolase Enzymes and Relatives  
Source URL: <http://bioweb.ensam.inra.fr/ESTHER/general?what=index>  
Document Type: **Repository** Times Cited: **2** (from All Databases)  
[  [View abstract](#) ]



- 1. Title: **Enzymatic Activity and Protein Interactions in Alpha/Beta Hydrolase Fold Proteins: Moonlighting Versus Promiscuity**  
Author(s): Marchot, Pascale; Chatonnet, Arnaud  
Source: PROTEIN AND PEPTIDE LETTERS Volume: **19** Issue: **2** Pages: **132-143** Published: **FEB 2012**  
Times Cited: **3** (from All Databases)  
 [  [View abstract](#) ]
- 2. Title: **ESTHER, the database of the alpha/beta-hydrolase fold superfamily of proteins**  
Author(s): Hotelier, T; Renault, L; Cousin, X; et al.  
Source: NUCLEIC ACIDS RESEARCH Volume: **32** Special Issue: **SI** Pages: **D145-D147** DOI: **10.1093/nar/gkh141** Published: **JAN 1 2004**  
Times Cited: **79** (from All Databases)  
  [  [View abstract](#) ]





## Principles

The Data Citation Principles cover purpose, function and attributes of citations. These principles recognize the dual necessity of creating citation practices that are both human understandable and machine-actionable.

These citation principles are not comprehensive recommendations for data stewardship. And, as practices vary across communities and technologies will evolve over time, we do not include recommendations for specific implementations, but encourage communities to develop practices and tools that embody these principles.

The principles are grouped so as to facilitate understanding, rather than according to any perceived criteria of importance.

### 1. Importance

Data should be considered legitimate, citable products of research. Data citations should be accorded the same importance in the scholarly record as citations of other research objects, such as publications<sup>[1]</sup>.

### 2. Credit and Attribution

Data citations should facilitate giving scholarly credit and normative and legal attribution to all contributors to the data, recognizing that a single style or mechanism of attribution may not be applicable to all data<sup>[2]</sup>.

### 3. Evidence

In scholarly literature, whenever and wherever a claim relies upon data, the corresponding data should be cited<sup>[3]</sup>.

### 4. Unique Identification

A data citation should include a persistent method for identification that is machine actionable, globally unique, and widely used by a community<sup>[4]</sup>.

### 5. Access

Data citations should facilitate access to the data themselves and to such associated metadata, documentation, code, and other materials, as are necessary for both humans and machines to make informed use of the referenced data<sup>[5]</sup>.

### 6. Persistence

Unique identifiers, and metadata describing the data, and its disposition, should persist -- even beyond the lifespan of the data they describe<sup>[6]</sup>.

### 7. Specificity and Verifiability

Data citations should facilitate identification of, access to, and verification of the specific data that support a claim. Citations or citation metadata should include information about provenance and fixity sufficient to facilitate verifying that the specific timeslice, version and/or granular portion of data retrieved subsequently is the same as was originally cited<sup>[7]</sup>.

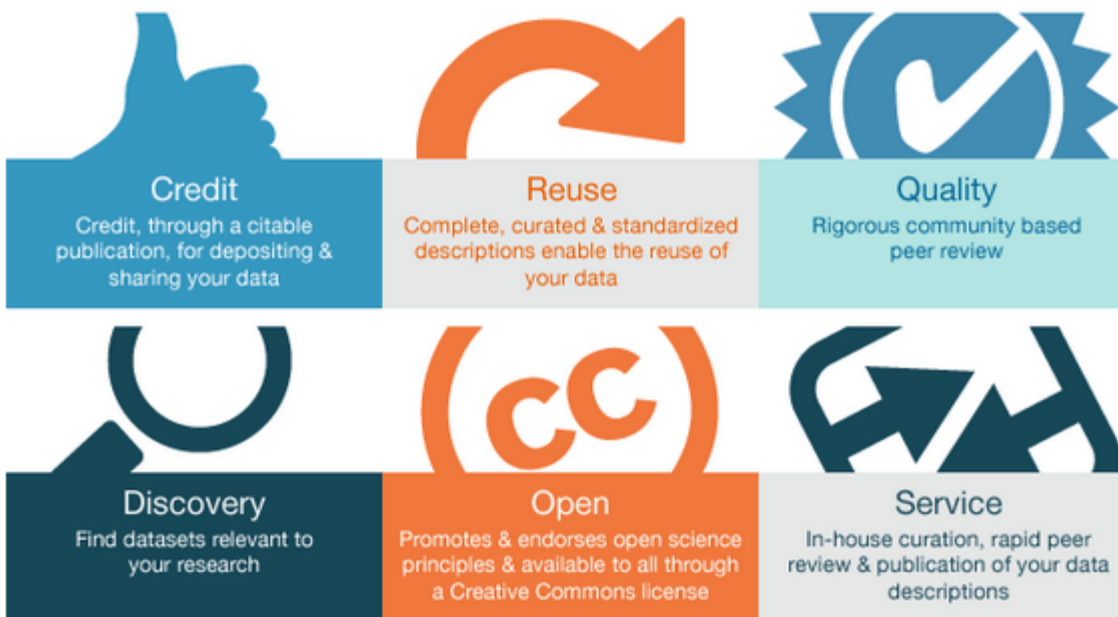
### 8. Interoperability and Flexibility

Data citation methods should be sufficiently flexible to accommodate the variant practices among communities, but should not differ so much that they compromise interoperability of data citation practices across communities<sup>[8]</sup>.



# Stratégie des grands éditeurs

## Helping you publish, discover, and reuse research data




### Welcome to *Scientific Data*

*Scientific Data* is a new open-access, online-only publication for descriptions of scientifically valuable datasets. It introduces a new type of content called the Data Descriptor, which will combine traditional narrative content with curated, structured descriptions of research data, including detailed methods and technical analyses supporting data quality. *Scientific Data* will initially focus on the life, biomedical and environmental science communities, but will be open to content from a wide range of scientific disciplines. Publications will be complementary to both traditional research journals and data repositories, and will be designed to foster data sharing and reuse, and ultimately to accelerate scientific discovery.


*Scientific Data* will launch in Spring 2014 and be open for submissions in Autumn 2013. Sign up for our

# Figshare – Digital Science (MacMillan/Nature)



My data    O. Hologne

My data 0% of private storage used

	Type <small>mouseover</small>	Date	Status	Statistics <small>public items only</small>
<input type="checkbox"/> esi18012012.xlsx	 DATASET	25.01.2012 13:31	<input type="button" value="DRAFT"/>	<input type="button" value="Add info"/>

showing all 1 items



Open Access

## Genomics Data

*Genomics Data* is an open access journal that considers articles on all aspects of genome-scale analysis:

1. Limited only to nucleic acids analysis
2. Microarray and Next-Gen Sequencing data
3. All organisms

Guide for Authors ▼

Submit Your Paper ▼

View Articles

## Elsevier Launches Genomics Data App in Collaboration with Illumina

Share This: [f](#) [t](#) [in](#) [g+](#)

*Genomics Data App now available in BaseSpace Apps*

**Cambridge, MA, May 20, 2013**

Elsevier, a world-leading provider of scientific, medical and technical products and services, announces the launch of the Genomics Data app in collaboration with Illumina, a sequencing and array-based solutions provider for genetics research.

The app allows researchers to share large data sets in Illumina's BaseSpace cloud, where they can easily be reviewed and included in Elsevier's open access journal, *Genomics Data*. The BaseSpace app serves as a gateway, opening data stored in the cloud to the publishing environment for expert peer review and wide dissemination in the form of a professional publication.

"In genomic studies we often generate large data sets of tremendous value. Without a means to publish these data sets and get credit for them, they easily get lost," said John Quackenbush, Co-Editor-in-Chief of *Genomics Data*. "By working together with Illumina to develop a reviewing pipeline from BaseSpace to *Genomics Data*, we've created a simple way for researchers to make genomic data accessible to publication in a journal."

Through BaseSpace, authors and reviewers will also have additional access to analytical tools, further enhancing *Genomics Data's* peer review process. This enhanced review option can give an article attached to the data set added credibility because reviewers are in the position to reproduce the analysis that the authors did.

"Uploading genomic data to BaseSpace has reached critical mass. The innovative development of the Genomics Data App demonstrates how Elsevier is able to create new and unique opportunities for researcher collaboration," said Alex Dickinson, Senior Vice President, Cloud Genomics at Illumina.

To view the app, please register for free on:

<https://basespace.illumina.com/apps/144144/Genomics-Data>

### Data Rescue Project Winner

[Nimbus Data Rescue Project Takes First Prize in Elsevier/IEDA Data Rescue Competition for Earth Sciences — oldWeather and Nuclear Explosion Signals Receive Honorable Mentions — see Press release](#)

Save time and more easily document and credit your research data using our expertise.

Research Data Services is a new group, exploring how Elsevier can help researchers share and annotate data. We are focused on getting data into discipline-specific repositories, bringing our core competencies in nomenclatures, informatics, repeatability and best practices in process, process management, impact assessment and analytics to overcome the barriers of research data sharing and curation.

### Principles

*Data must be open — openly available, supporting open science.*

*Anything we do must be done with and through the research community and funding agencies.*

### Goals

*Help increase archiving and sharing of research data.*

*Help increase the discoverability and value of shared data with annotation, context, provenance.*

*Foster and assist with the credit and impact assessment of research data for the researcher, the institution, and the funding bodies, making research data first-class research outputs.*

### Challenges





# Perspectives et conclusions



# Repenser le système

<http://dx.doi.org/10.1038/ng0411-281>

Figure 2: A proposal for the future of scholarly communication.

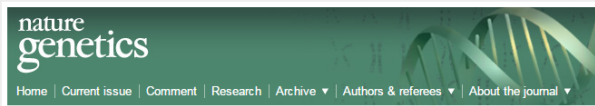
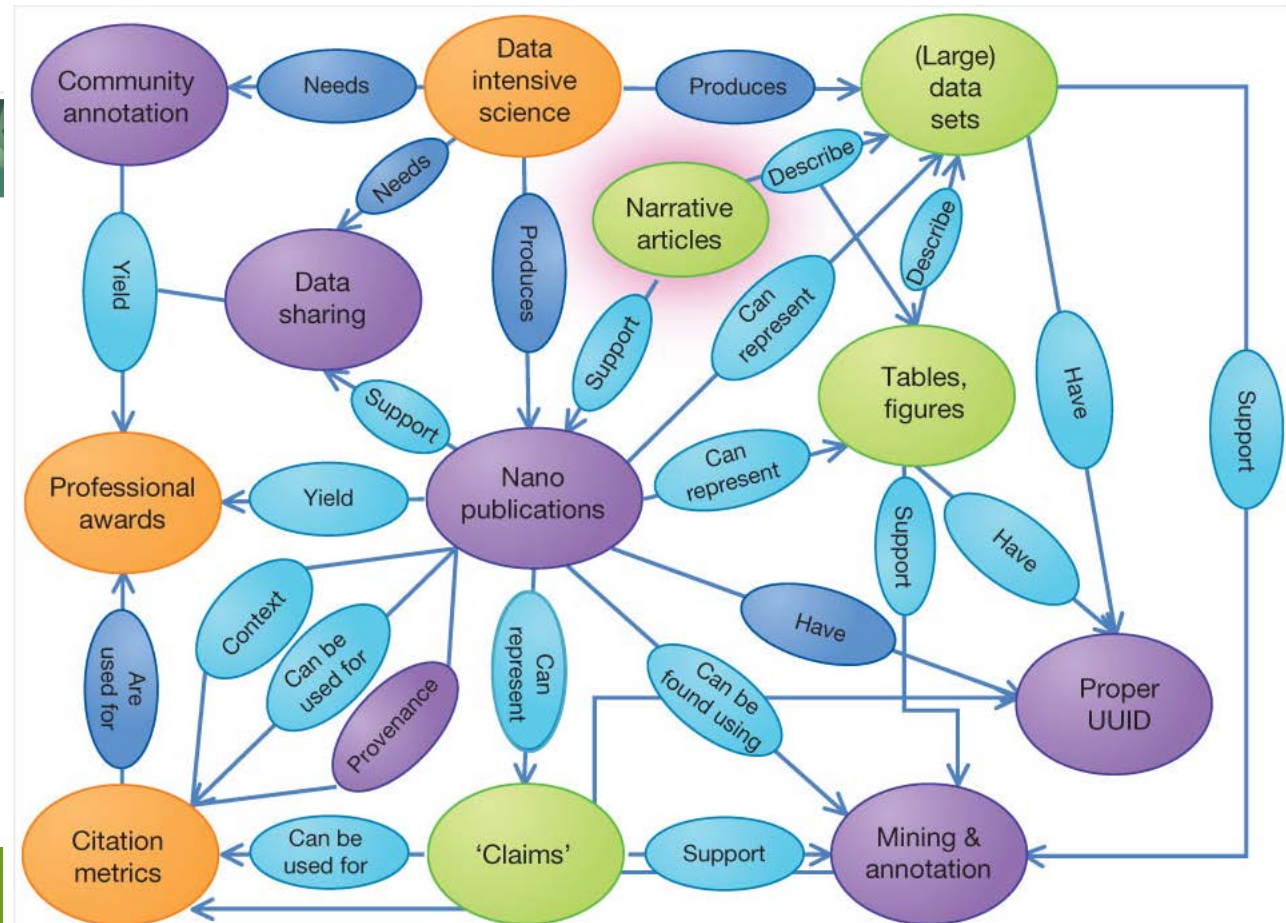
From

The value of data

Barend Mons, Herman van Haagen, Christine Chichester, Peter-Bram 't Hoen, Johan T den Dunnen, Gertjan van Ommen, Erik van Mulligen, Bharat Singh, Rob Hooft, Marco Roos, Joel Hammond, Bruce Kiesel, Belinda Giardine, Jan Velterop, Paul Groth & Erik Schultes

*Nature Genetics* 43, 281–283 (2011) | doi:10.1038/ng0411-281

Published online 29 March 2011



home > archive > issue > commentary > full text

NATURE GENETICS | COMMENTARY

## The value of data

Barend Mons, Herman van Haagen, Christine Chichester, Peter-Bram 't Hoen, Johan T den Dunnen, Gertjan van Ommen, Erik van Mulligen, Bharat Singh, Rob Hooft, Marco Roos, Joel Hammond, Bruce Kiesel, Belinda Giardine, Jan Velterop, Paul Groth & Erik Schultes

Affiliations | Contributions | Corresponding author

*Nature Genetics* 43, 281–283 (2011) | doi:10.1038/ng0411-281  
Published online 29 March 2011



# Conclusion

- ❖ Les données, nouvel eldorado des éditeurs ?
- ❖ Evolution de l'objet « article » : texte, données, traitement/code
- ❖ Interconnexion des informations
  - ✓ linked data



# Merci pour votre attention...

[odile.hologne@versailles.inra.fr](mailto:odile.hologne@versailles.inra.fr)

@Holo\_08



# Lectures et sites intéressants

Research Data Alliance <http://rd-alliance.org/>

Science as an open entreprise <http://royalsociety.org/policy/projects/science-public-enterprise/report/>

JISC <http://www.jisc.ac.uk/whatwedo/programmes/mrd.aspx>

Digital curation center

<http://www.dcc.ac.uk/>

<http://www.dcc.ac.uk/training/rdm-librarians> (formation pour les « Librarian »)

<http://www.dcc.ac.uk/events/idcc13/programme-presentations> Conf #idcc13

Australian national data service <http://www.ands.org.au/>

LERU Roadmap for research data

[http://www.uzh.ch/research/LERU\\_Roadmap\\_for\\_Research\\_data.pdf](http://www.uzh.ch/research/LERU_Roadmap_for_Research_data.pdf)

LIBER <http://liber.library.uu.nl/index.php/lq/article/view/9173/9922>