

# Scientific Data in short



## CREDIT

引用可能な形で研究データを公開できクレジットが得られる



## REUSE

データ記述をキュレーションによって標準化し再利用に役立てます



## QUALITY

各コミュニティの基準に基づいた厳格な査読でクオリティを保証



## SEARCH

ユーザーの研究に関係のあるデータセットを発見できます



## OPEN SCIENCE

クリエイティブコモンズライセンスでデータの再利用を推進



## SERVICE

迅速な査読、出版と社内キュレーション（データ自体も査読）

# *Nature* and NPG data policies

---

- Enforce community database deposition
- Encourage community database development
- Launch *Scientific Data*
- Nature-journal editors encourage submissions of Data Descriptors to *Scientific Data*

Website for authors & referees on data policies

[www.nature.com/authors/policies/availability](http://www.nature.com/authors/policies/availability)

# Open data to enable publication of more reproducible and reliable research

Nov  
2014

## ANNOUNCEMENT

### Data-access practices strengthened

In our continued drive for reproducibility, *Nature* and the *Nature* research journals are strengthening our editorial links with the journal *Scientific Data* and enhancing our data-availability practices. We believe that this initiative will improve support for authors looking for appropriate public repositories for their research data, and will increase the availability of information needed for the reuse and validation of those data.

In 2013, *Nature* journals introduced new editorial measures to promote reproducibility, and we continue to evaluate their impact and refine our policies. Our newly strengthened data-availability practices ([go.nature.com/05ykhe](http://go.nature.com/05ykhe)) reflect our preference that data be deposited in public repositories, and encourage researchers to expand on work published in the *Nature* journals by publishing further information in *Scientific Data*.

Community-supported, specialized data repositories are usually the best way to share large data sets. General, unstructured repositories, such as figshare and Dryad, provide options where no community repository exists, and are preferable to publishing data as Supplementary Information. Supplementary materials have size limitations and do not always provide optimal file and viewing formats, particularly for large and complex data sets. But where no repository — or publication focused on detailed descriptions of data sets — exists, supplementary materials have often been the best option.

*Scientific Data* ([go.nature.com/iyu9qh](http://go.nature.com/iyu9qh)), which launched this year, offers authors another way to maximize the value of their data sets for further research — for themselves and for the scientific community.

Its primary article type, the Data Descriptor, provides more detail to improve the data's discoverability, interpretability and

reusability — as well as allowing the highest credit to be given to the authors who created the data set.

We are now rolling out a new process under which, when they accept a manuscript containing appropriate data sets, editors of *Nature* and *Nature* research journals will encourage authors to submit the data sets to *Scientific Data* as a Data Descriptor ([go.nature.com/utfvfo](http://go.nature.com/utfvfo)).

Authors may also submit a Data Descriptor manuscript alongside a manuscript for a *Nature* journal. If appropriate, they could publish the descriptor first, without compromising the novelty of future primary-research articles based on the data. In these cases, authors are encouraged to consult with the editor of their target journal to ensure that prior publication of a Data Descriptor is acceptable. (Note that other publishers may have different policies.)

*Scientific Data's* peer-review and in-house curation processes focus on ease of reuse. A data-curation editor reviews data files, checks their format, archiving and annotations, and works with authors to produce a standardized, machine-readable summary of the study in the ISA-Tab format (S. Sansone *et al. Nature Genet.* **44**, 121–126; 2012).

Data Descriptors can accommodate all data types, including raw data and updated data sets generated after initial publication. They can also show the controls required for validation of the data set, which may have been excluded from the primary paper because of space limitations. *Scientific Data's* editorial process assesses repositories and helps to ensure that data are placed in the correct one. *Nature's* enhanced data-availability policy now directs authors to a list of approved repositories ([go.nature.com/jpm768](http://go.nature.com/jpm768)).

Several articles published in *Nature* research journals already have complementary articles in *Scientific Data* (such as A. Baud *et al. Sci. Data* **1**, 140011 (2014) and F. Roquet *et al. Sci. Data* **1**, 140028; 2014). As science evolves and produces ever-increasing amounts of data, those data must be collected, organized, curated, quality-checked and made available on the right platform so that they can be easily discovered and reused. Stronger links with *Scientific Data* and our data-availability practices aim to achieve this. ■

# Scientific Data – recommended data repositories

## Life Sciences

### **Nucleic acid sequence**

DDBJ/ ENA/ GenBank/ dbSNP/ EVA/ dbVar  
DGVa/ EBI Metagenomics / NCBI Trace  
Archive/ NCBI SRA

### **Protein sequence**

Uniprot

### **Molecular & supramolecular structure**

PCDDDB/ COD/ CXIDB/ BMRB/ EMDB/ wwPDB

### **Neuroscience**

NeuroMorpho.org/ FCP/INDI/ OpenfMRI

## **Omics**

**Functional Genomics:** ArrayExpress/ GEO/  
GenomeRNAi/ dbGAP/ EGA/ DIP/ IntAct

**Metabolomics:** MetaboLights

**Proteomics:** PeptideAtlas/ PRIDE/  
ProteomeXchange

### **Taxonomy & species diversity**

ITIS/ KNB/ NCBI Taxonomy/ GBIF

## **Life-science community resources**

EuPathDB/ FlyBase/ Influenza Research Database/  
MGI/ NAHDAP/ NDAR/ RGD/ VectorBase/ Xenbase/  
ZFIN/ BioGRID/ Cancer Imaging Archive/  
ClinicalTrials.gov/ SICAS medical image repository/  
PhysioNet/

## **Chemistry & chemical biology**

caNanoLab/ KiMoSys/ PubChem

## **Environmental & geoscience**

NCDC/ OPNL DAAC/ CARD/ WDCC/ NERC Data  
Centres/ LTER Community Data Repository/  
PANGAEA/ AADC

## **Physics, astrophysics & astronomy**

Reaction Database Standard Search Interface/  
SIMBAD Astronomical Database/ UK Solar System  
Data Centre

## **Social science**

Harvard Dataverse Network

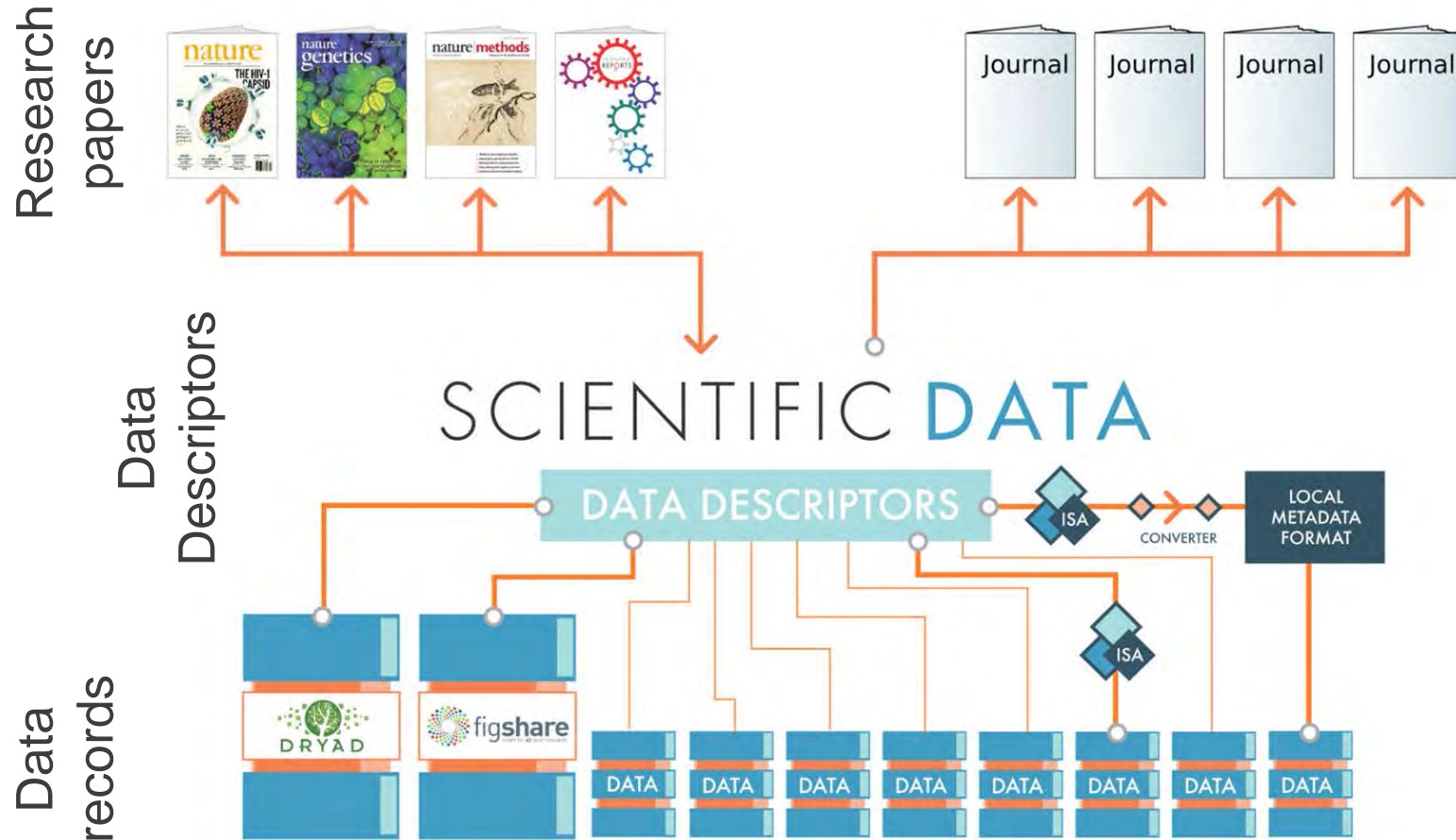


Science and Education

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

Data Policies > Recommended Repositories

# A value added component in a growing data publishing ecosystem



# Introducing Scientific Data

Initiative for Open  
Scientific Data at  
Nature Publishing  
Group :

Yoko Shintani  
Open Research Marketing  
Manager, Nature Publishing  
Group

[y.shintani@nature.com](mailto:y.shintani@nature.com)