

Sharing and publishing your data

PUBLISHERS CARE & ARE HERE FOR YOU

Magdalena Skipper

Editor-in-Chief, *Nature*

14.11.2017

Something you already know because you're here...

Data sharing encourages more connection and collaboration between researchers. In a time of reduced investment for science and research, data sharing is more efficient because it allows researchers to share resources.

Benefits of data sharing

1. Transparency and reproducibility

Lack of underlying data availability makes it difficult to verify and replicate results. If data were always be made available with publications, transparency of research would be greatly improved

2. Reuse

The availability of raw data allows other researchers to reuse the data. Not just for replication purposes, but to pursue new research questions

3. Data quality

The need to make data visible to others can lead to better data management and quality

4. Credit

Citing previously used data forms the basis for a data credit system. Data citation is the first step towards incentivizing researchers to share their data and to measure their impact

5. Informed decision-making, policy and public perception of science

Data citation (with a DOI) makes ...

- data visible
- accessible for re-use
- reproducible for transparency
- It easier to measure data impact
- enhances dividends from the original investment by funders and scientists

At the moment, only about 10 percent of authors cite or acknowledge the datasets used in research papers properly.



The State of Open Data 2018

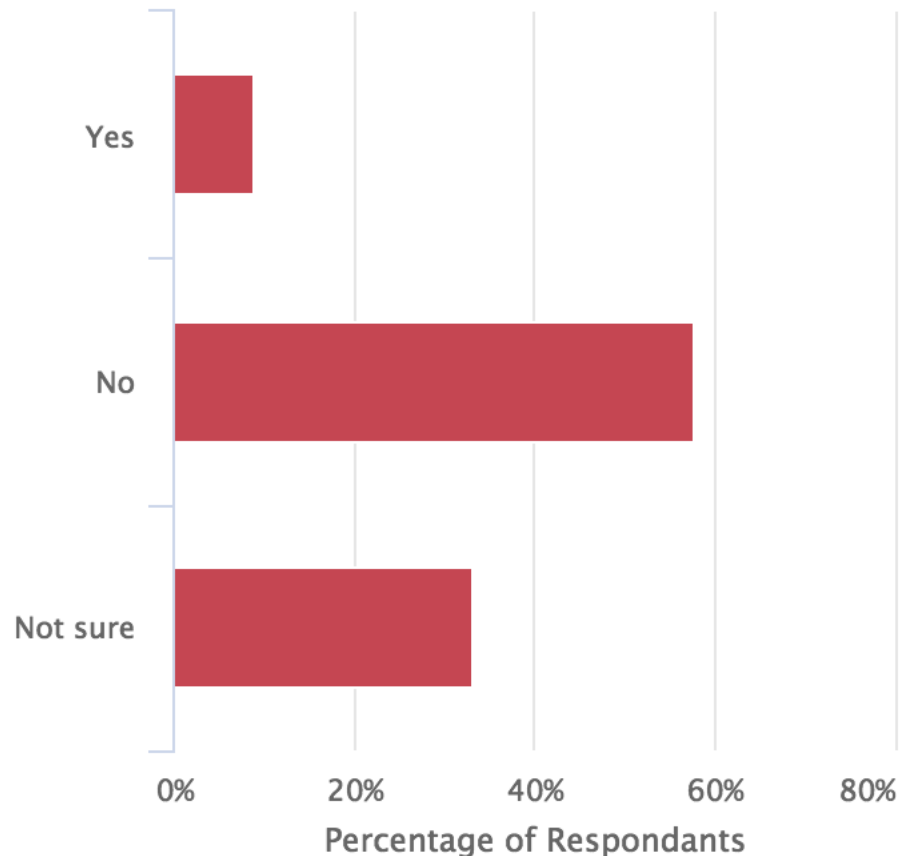
Global Attitudes towards Open Data

#Stateofopendata

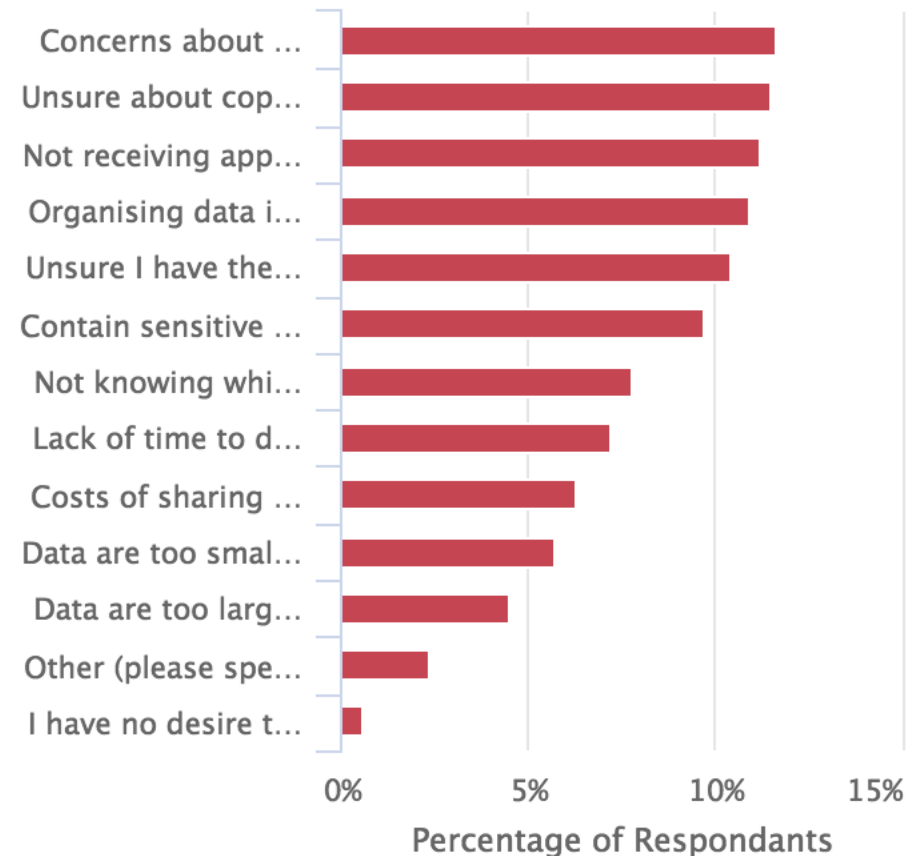
Figshare (Digital Science)

<https://www.digital-science.com/blog/news/the-state-of-open-data-2018-global-attitudes-towards-open-data-stateofopendata/>

Do researchers get sufficient credit for sharing data?

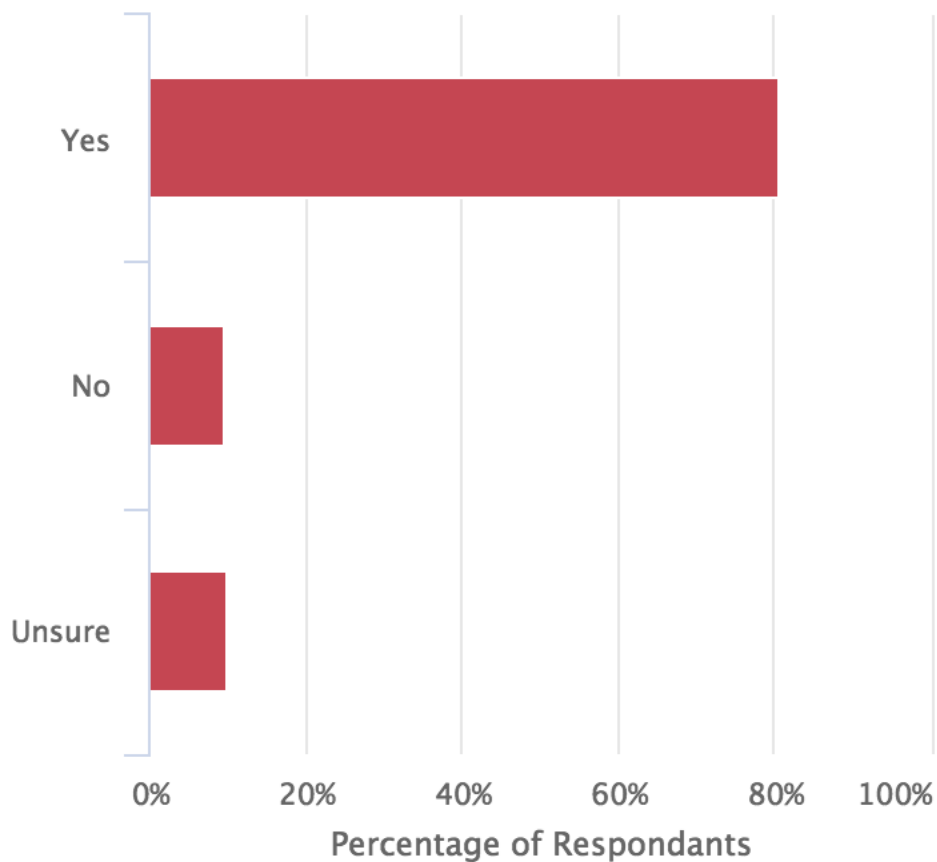


What concerns do you have sharing dataset?

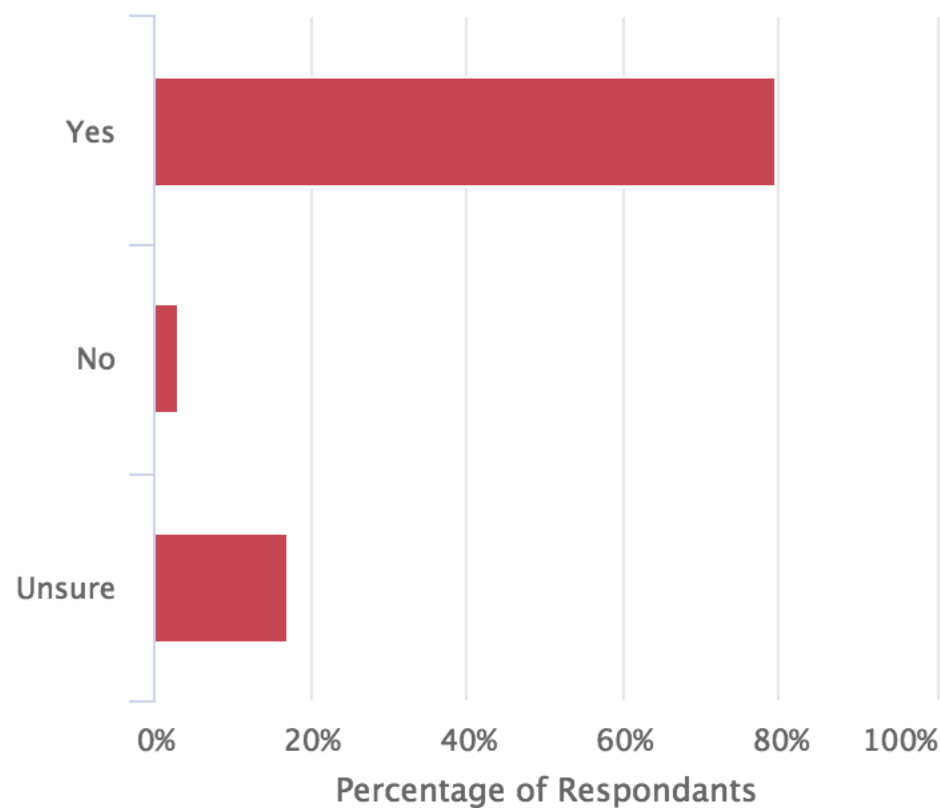


Researchers clearly do not feel there is sufficient credit given for sharing data. When asked directly what concerns they have about sharing, they overwhelmingly skip the question. We conclude that lack of knowledge about credit mechanisms is responsible for this.

Are you aware of any research data that is openly available?



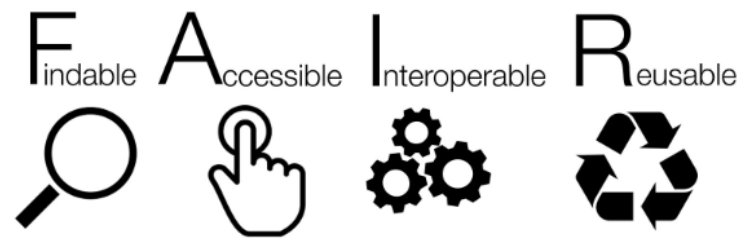
Would you be willing to reuse open data for your own research in the future?



Researchers are broadly aware of existing open data and are willing to reuse other's data in their own work.

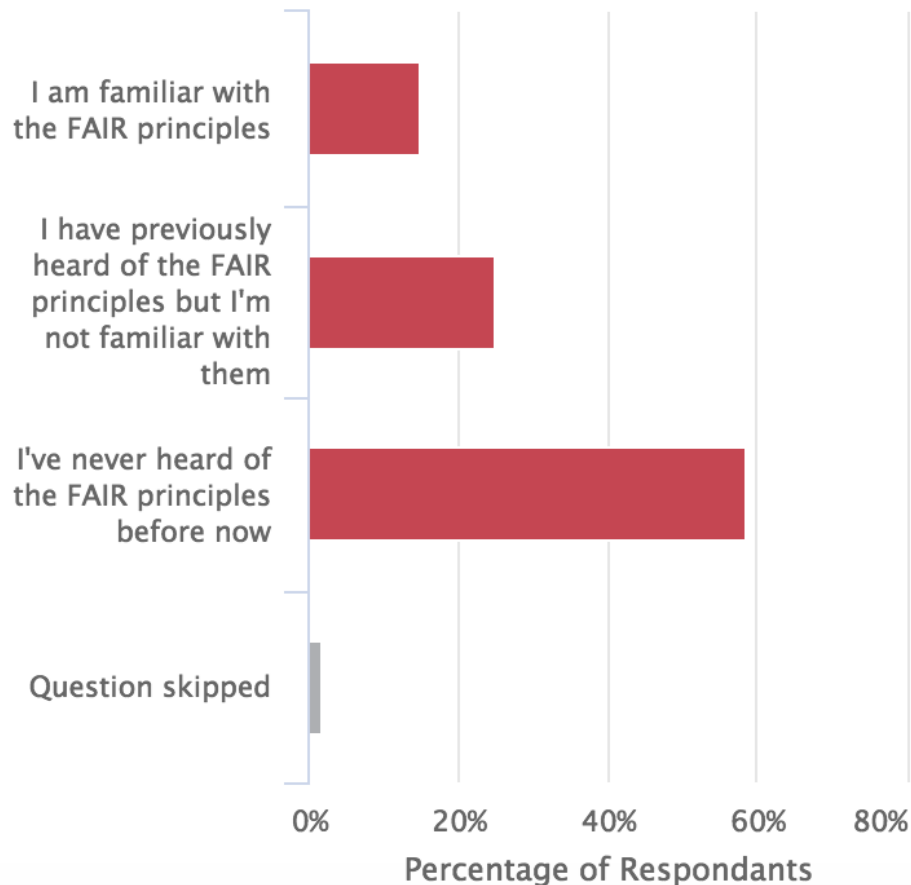
The FAIR data principles are a fairly new development in Open Data introduced in 2016 to give researchers measurable goals when sharing data. Learn more about FAIR here [DOI:10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)

Over 55% of survey respondents had never heard about FAIR before taking this survey, we will track the exposure of FAIR principles in future surveys.

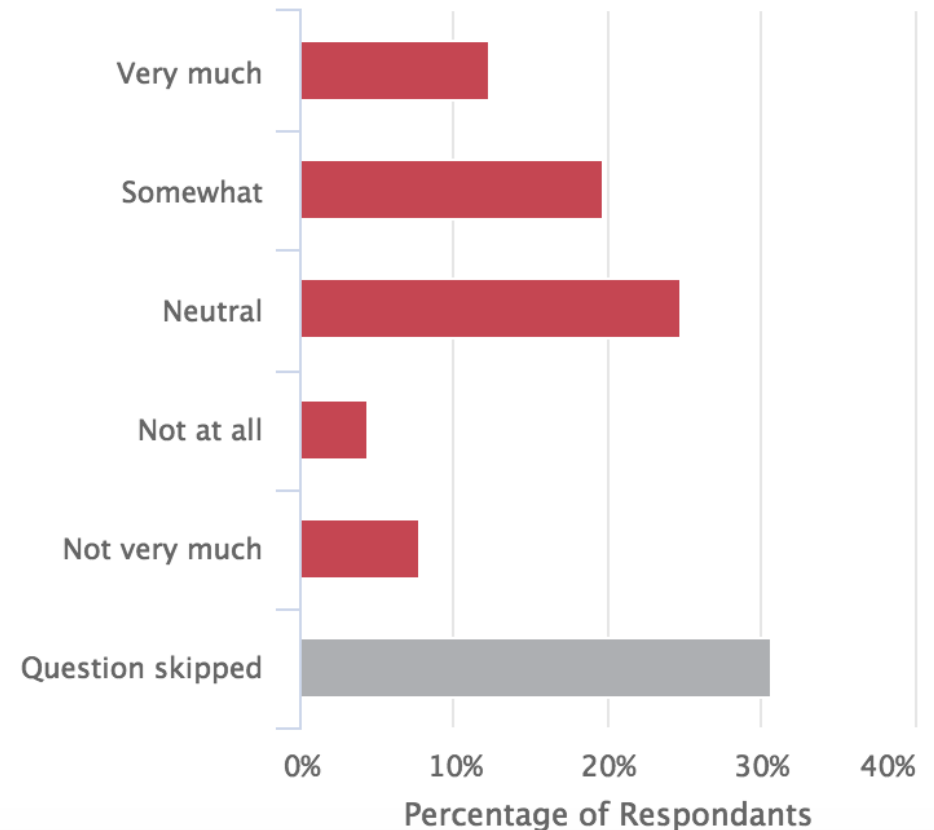


Source: [Wikimedia](#).

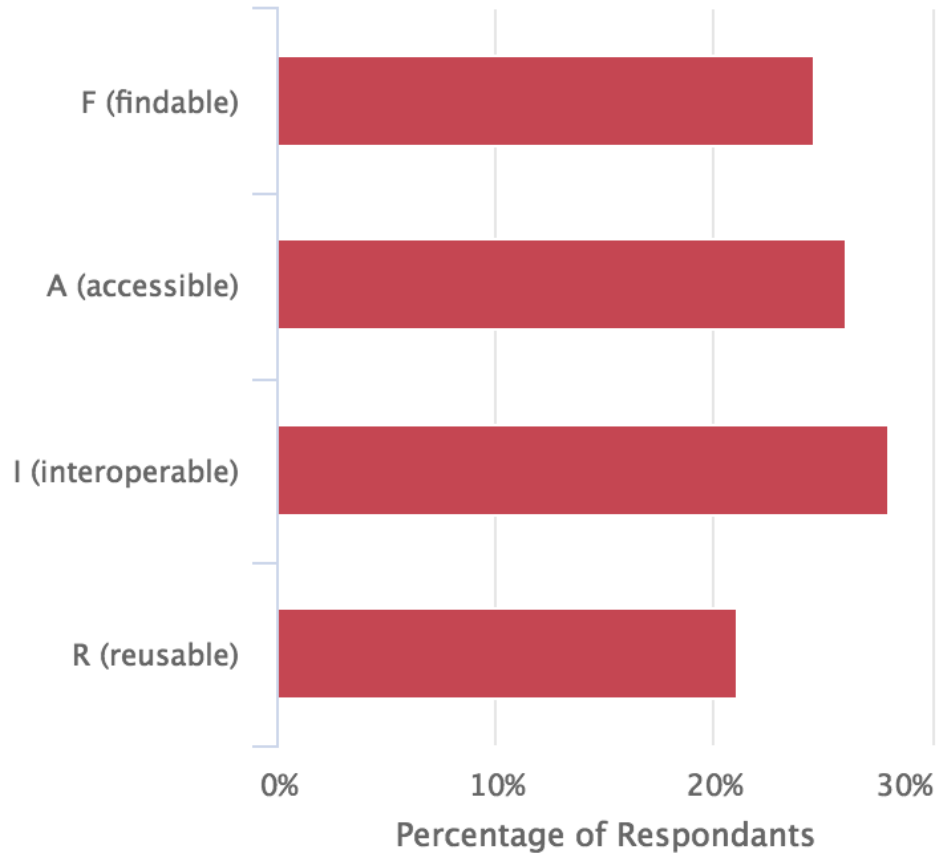
What's your familiarity with FAIR?



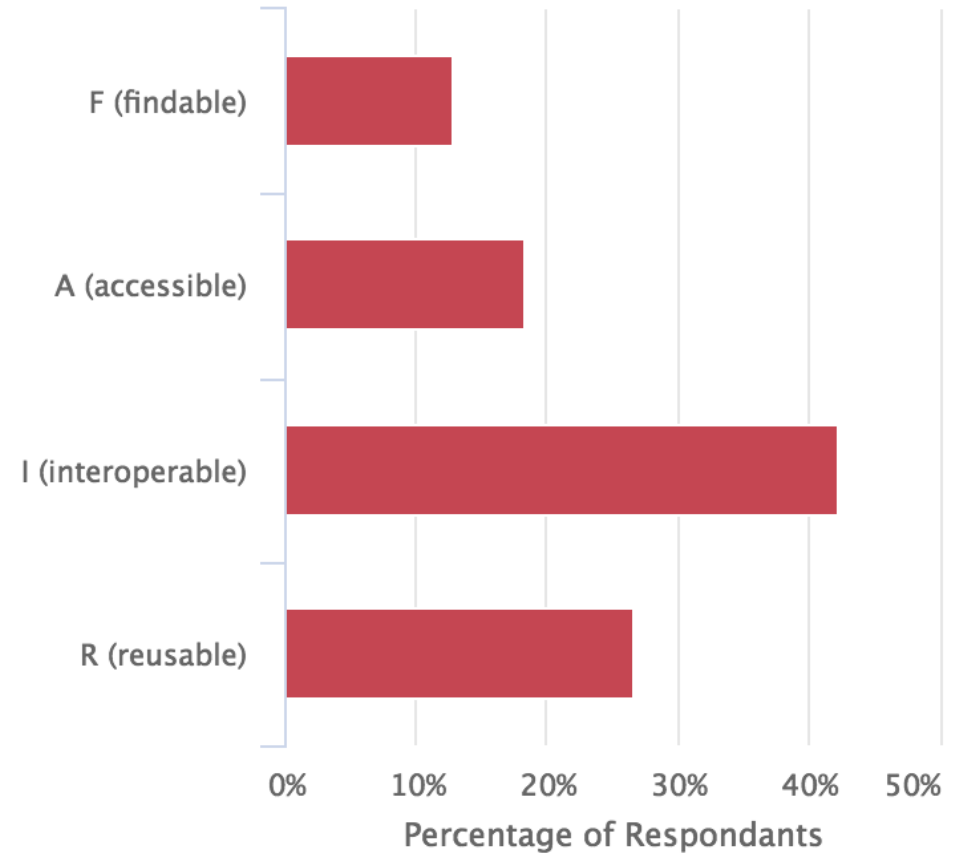
How compliant is your data with FAIR?



Which principle of FAIR is your data least compliant with

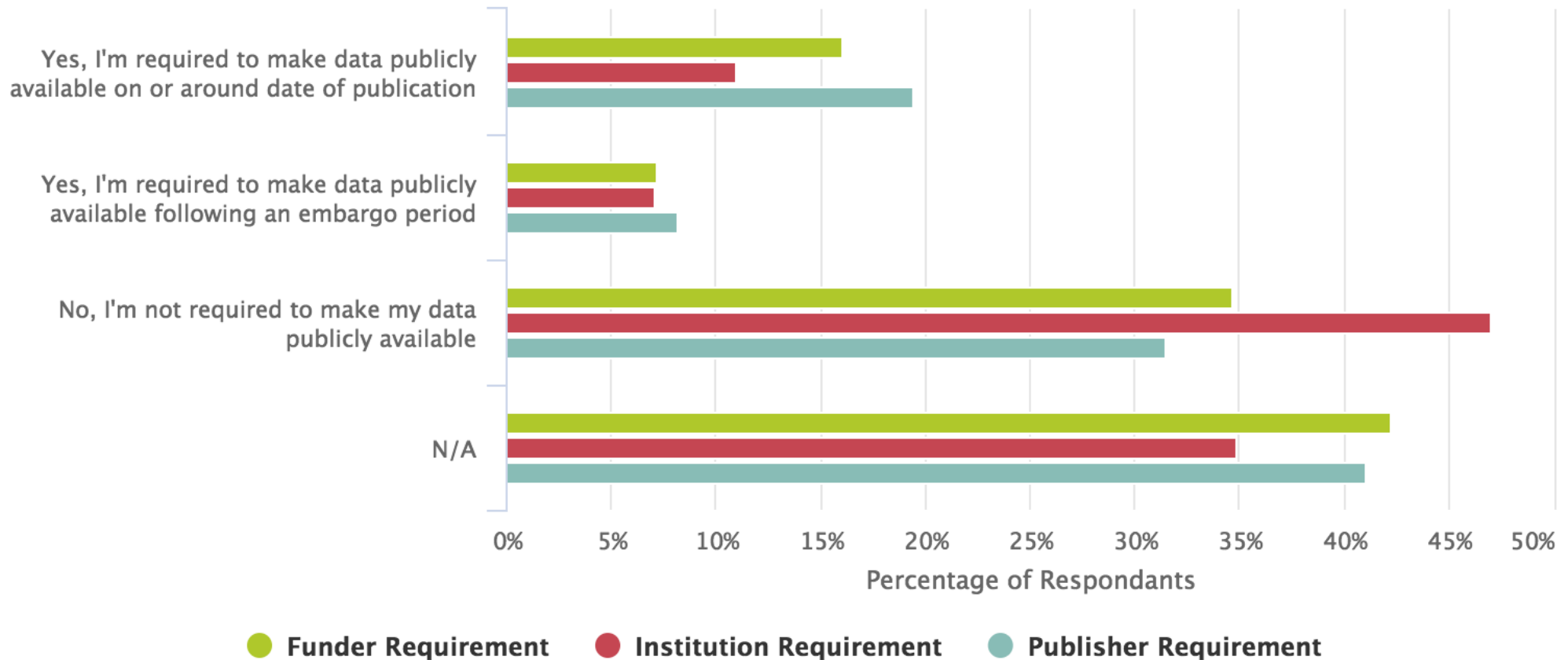


Which FAIR principle do you think needs better definition?



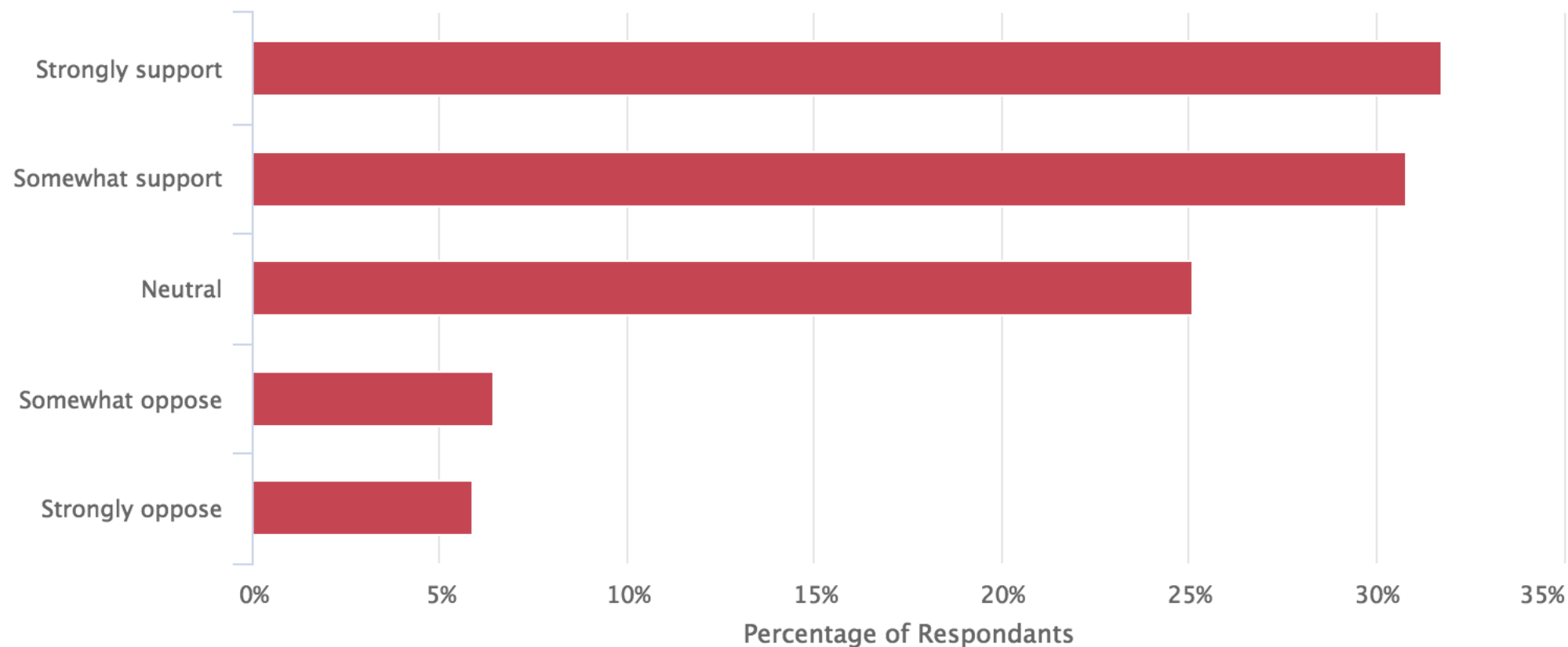
Interoperability is clearly the least understood FAIR principle.

For your most recent piece of research were you subject to any data sharing policies?



How supportive would you be of a national mandate for making primary research data openly available?

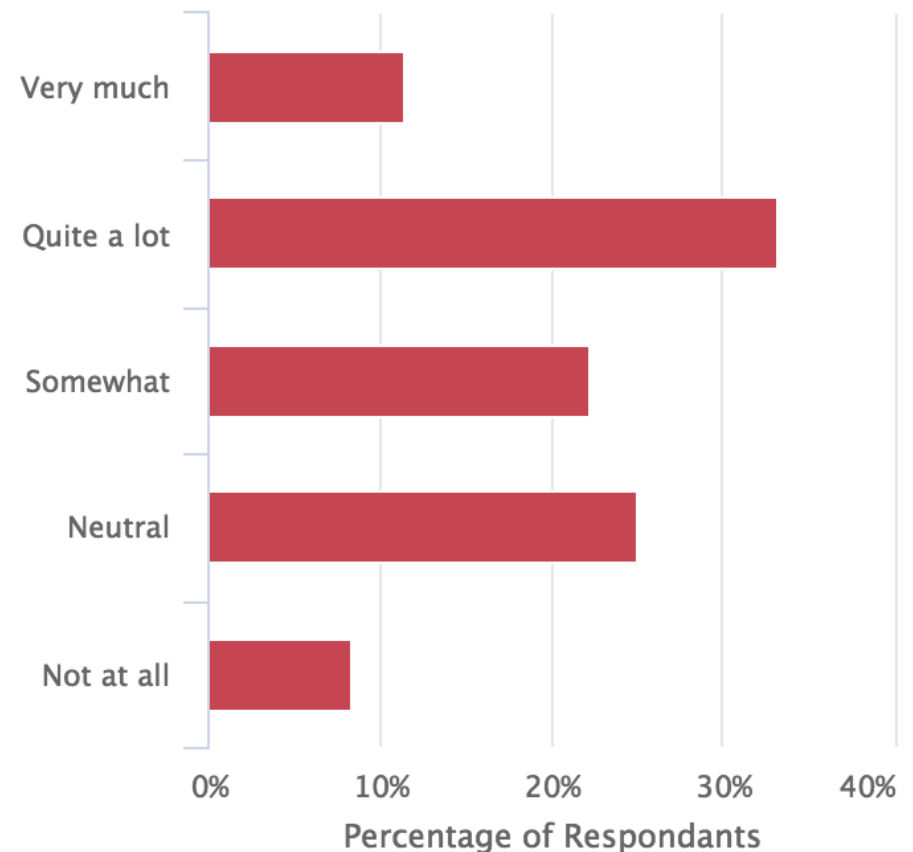
(By open, we mean free to access, reuse, repurpose, and redistribute)



What most motivates you to share data?

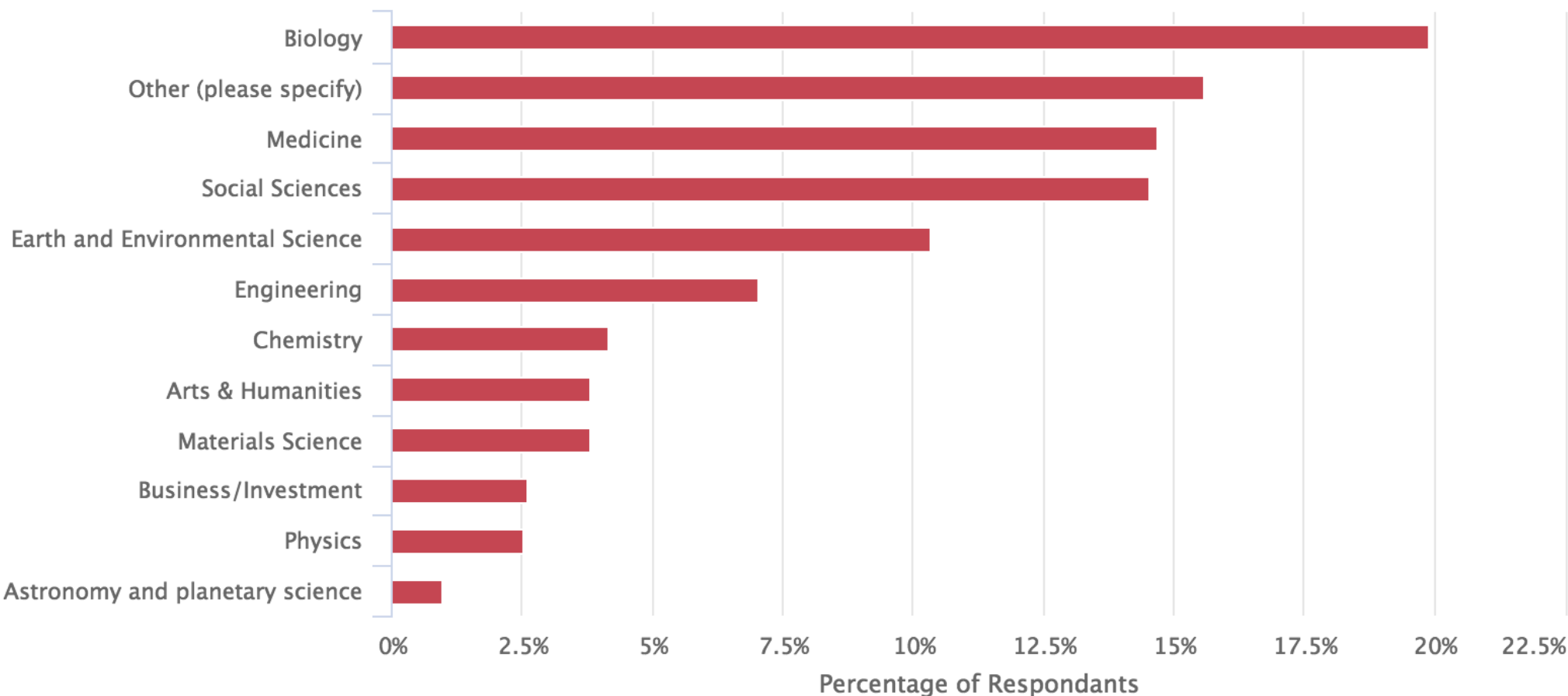


How much would data citations motivate you?



Researchers are primarily motivated to share data for increased research impact and other altruistic sentiments; independent of whether respondents are only allowed to select one or many motivators.

Which of the following best describes your area of interest?



Over 15% of respondents provided additional details about their research interests, indicating these narrow categories are of less use in today's interdisciplinary research space.

Understanding researcher behaviours globally

Practical challenges for researchers in data sharing



Stuart, David; Baynes, Grace; Hrynaszkiewicz, Iain; Allin, Katie; Penny, Dan; Lucraft, Mithu; Astell, Mathias (2018):
Whitepaper: Practical challenges for researchers in data sharing
<https://doi.org/10.6084/m9.figshare.5975011.v1>

Importance of data discoverability

76%
of respondents

highly rate the importance of their data being discoverable:
most popular ranking was 10/10

Importance
out of 10

7.7 South America

7.6 Asia

7.3 Europe

7.2 North America

6.9 Australasia

Importance
out of 10

6.5

Other
sciences

6.6

Physical
sciences

7.2

Medical
sciences

7.7

Earth
sciences

7.8

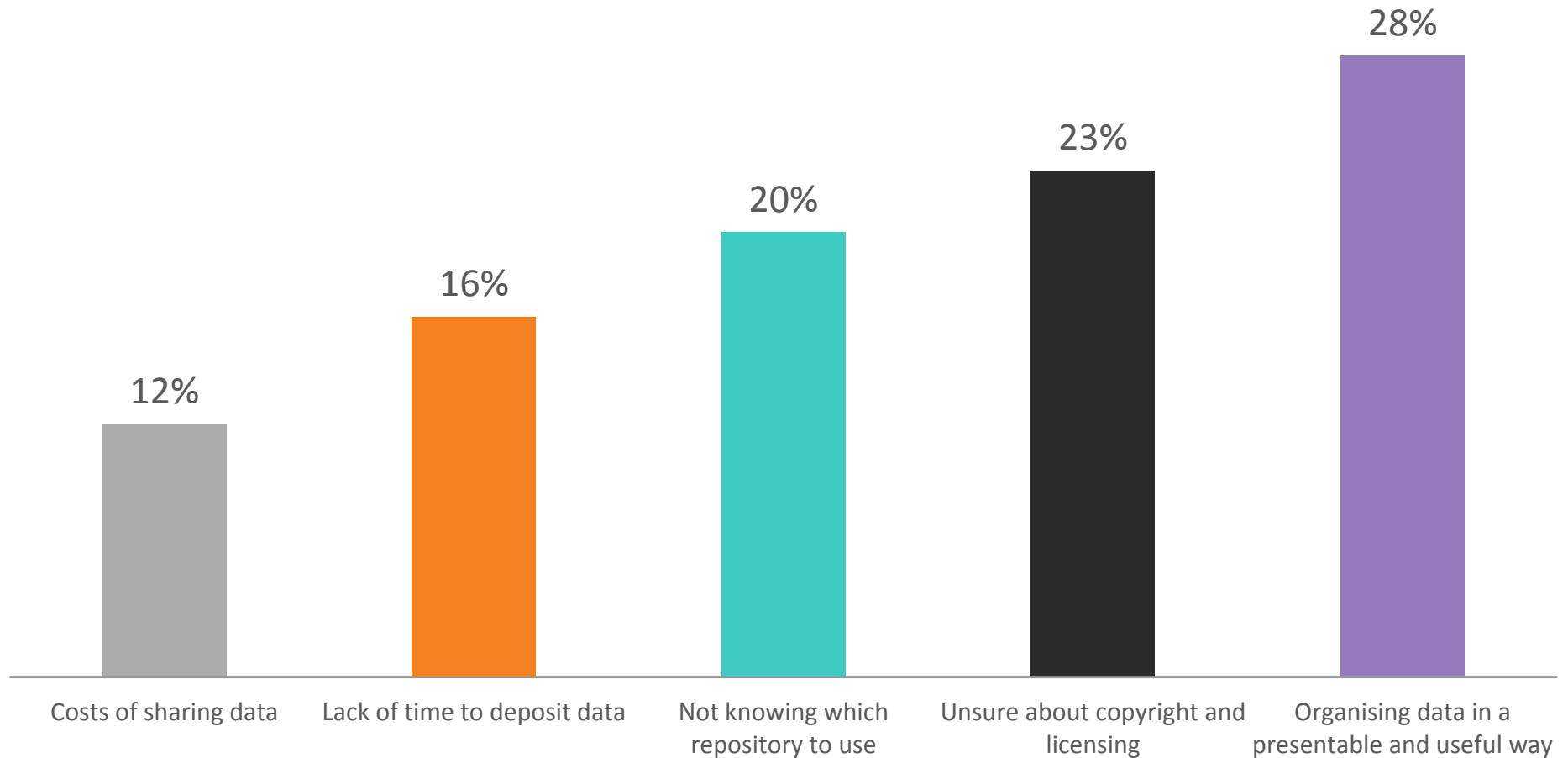
Biological
sciences



Majority of researchers share their data in some way

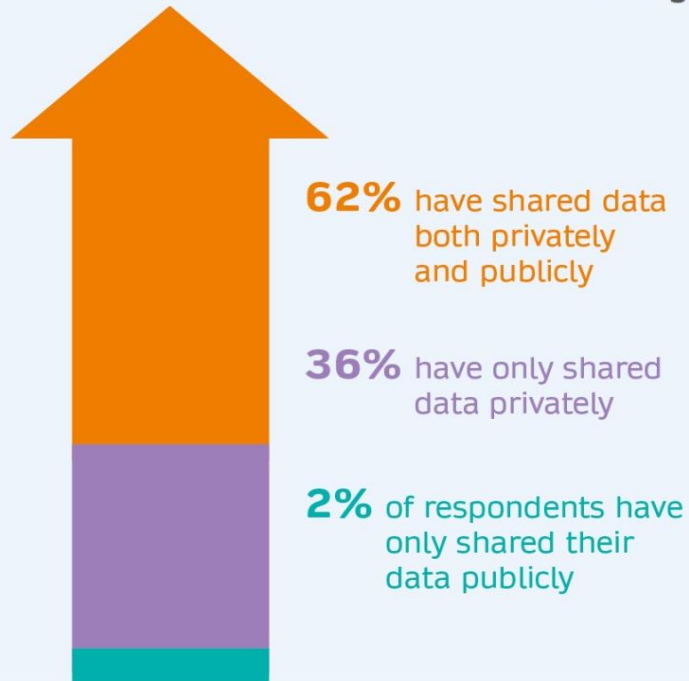
63% of respondents stated that they generally **submitted data files as supplementary information**, deposited the files in a **repository**, or both

Problems researchers have in sharing data by percentage



From a Springer Nature researcher survey (n=7719). <https://doi.org/10.6084/m9.figshare.5975011>

Private sharing of data is a lot more common than public sharing of data



The three most common methods of private sharing were:



email (65%)



USB or flash drives (41%)



file sharing services (39%)

The three most common ways of public sharing were: (n=569)



supplementary information to journal articles (51%)

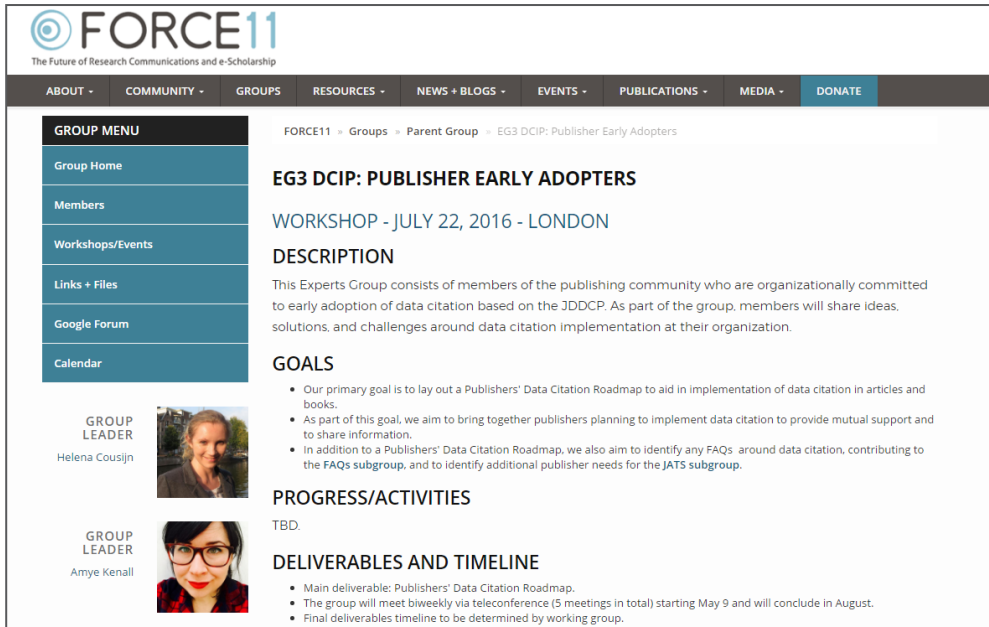


lab or personal website (27%)



subject specific repository or data archive (25%)

Collaboration to make things work at scale



FORCE11
The Future of Research Communications and e-Scholarship

ABOUT - COMMUNITY - GROUPS - RESOURCES - NEWS + BLOGS - EVENTS - PUBLICATIONS - MEDIA - DONATE

GROUP MENU

- Group Home
- Members
- Workshops/Events
- Links + Files
- Google Forum
- Calendar

FORCE11 » Groups » Parent Group » EG3 DCIP: Publisher Early Adopters

EG3 DCIP: PUBLISHER EARLY ADOPTERS

WORKSHOP - JULY 22, 2016 - LONDON

DESCRIPTION

This Experts Group consists of members of the publishing community who are organizationally committed to early adoption of data citation based on the JDDCP. As part of the group, members will share ideas, solutions, and challenges around data citation implementation at their organization.

GOALS

- Our primary goal is to lay out a Publishers' Data Citation Roadmap to aid in implementation of data citation in articles and books.
- As part of this goal, we aim to bring together publishers planning to implement data citation to provide mutual support and to share information.
- In addition to a Publishers' Data Citation Roadmap, we also aim to identify any FAQs around data citation, contributing to the FAQs subgroup, and to identify additional publisher needs for the JATS subgroup.

PROGRESS/ACTIVITIES

TBD.

DELIVERABLES AND TIMELINE

- Main deliverable: Publishers' Data Citation Roadmap.
- The group will meet biweekly via teleconference (5 meetings in total) starting May 9 and will conclude in August.
- Final deliverables timeline to be determined by working group.

GROUP LEADER
Helena Cousijn

GROUP LEADER
Amye Kenall

<https://www.force11.org/group/dcip/eg3publisherearlyadopters>

- Working group to define a Publishers' Data Citation Roadmap to aid in implementation of data citation in articles and books
- Covers policy, author guidelines and content production and technical issues
- Many publishers collaborating with infrastructure providers and policy makers to take a common approach
- The process should be simple for researchers while delivering benefits of better measurement of credit and reuse

Slide c/o Iain Hrynaszkiewicz

A Data Citation Roadmap for Scientific Publishers. Helena Cousijn, Amye Kenall, Emma Ganley, Melissa Harrison, David Kernohan, Fiona Murphy, Patrick Polischuk, Maryann Martone, Timothy Clark
 bioRxiv 100784; doi: <https://doi.org/10.1101/100784>

This article presents a practical roadmap for scholarly publishers to implement data citation in accordance with the Joint Declaration of Data Citation Principles (JDDCP), a synopsis and harmonization of the recommendations of major science policy bodies.



New Results

A Data Citation Roadmap for Scientific Publishers

 Helena Cousijn,  Amye Kenall,  Emma Ganley,  Melissa Harrison,  David Kernohan,  Thomas Lemberger,  Fiona Murphy,  Patrick Polischuk, Simone Taylor,  Maryann Martone,  Timothy Clark

doi: <https://doi.org/10.1101/100784>

This article is a preprint and has not been peer-reviewed [what does this mean?].

Why and How to Cite Data – For Authors

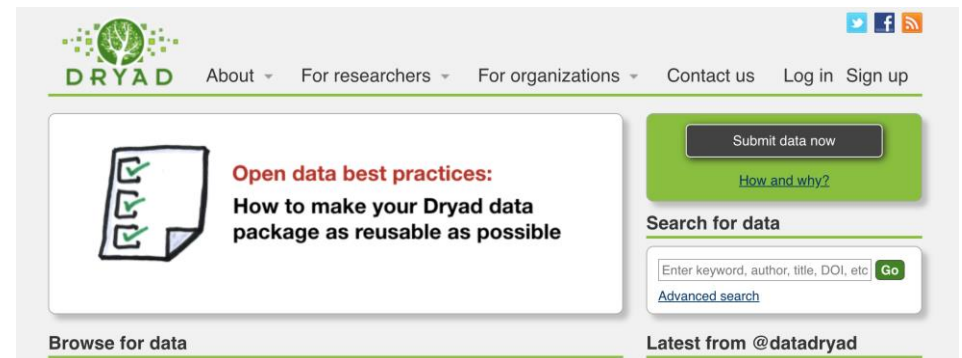
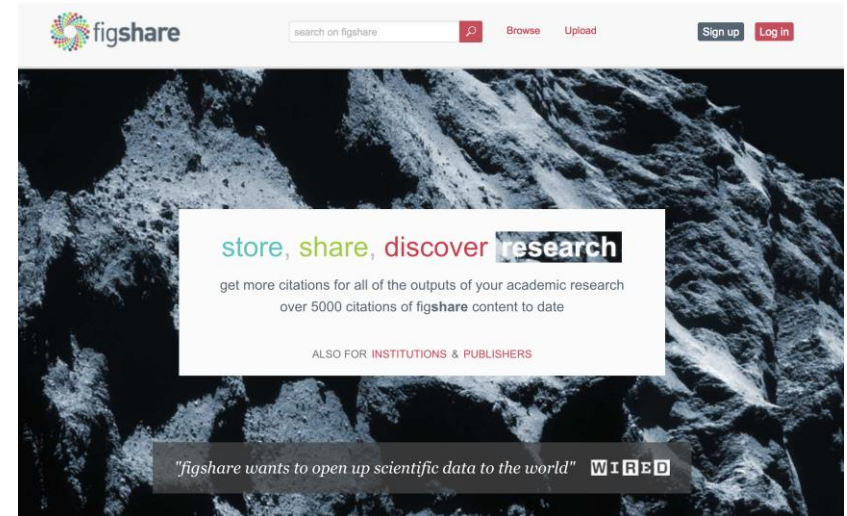
<i>Why?</i>	<ol style="list-style-type: none"> 1. Supports reproducibility and validation of results, allows data reuse. 2. Provides credit for data generators. 3. Publications linked to publicly available data have been associated with increased citations. 4. Improves connectivity and provenance tracking of data described in publications. 5. Data sharing is increasingly required by funders, publishers, and institutions.
<i>How?</i>	<ol style="list-style-type: none"> 1. For primary data: determine an appropriate long-term repository for data archiving. Your publisher should provide access to a list of acceptable archival repositories. 2. Deposit your data and get an accession number or dataset DOI from the repository. 3. For secondary data: cite what you use. When using the data of others, cite both the related peer-reviewed literature and the actual datasets used. 4. Include ‘formal’ data citations whenever possible: When datasets have formal, stable identifiers or accession numbers, they should be included in the main reference list. 5. Be as complete as possible, but don’t invent metadata: If a data record does not have a clear author/creator or title, don’t make one up. 6. Refer to your publisher’s <i>Guide for Authors</i> to format your dataset reference.

Why and How to Cite Data – For Publishers

<i>Why?</i>	<ol style="list-style-type: none"> 1. Help authors and journals to comply easily with funder mandates. 2. Improve author service by simplifying policies and procedures and increasing the visibility and connectivity of their articles and data. 3. Improve editor and peer reviewer service with better guidelines and support for data and visibility of data in the peer review process. 4. Improve reader and author service with more consistent links to data. 5. Support editorial goals to publish more open and reproducible research. 6. Make the most of your repository partnerships.
<i>How?</i>	<ol style="list-style-type: none"> 1. Revise editor training and advocacy material 2. Revise reviewer training material 3. Update information for authors by: <ol style="list-style-type: none"> a. providing guidance on author responsibilities and a policy on data citation; b. asking authors to provide a Data Availability Statement; c. specifying how to format data citations; and d. providing detailed guidance on suitable repositories. 4. Update guidelines for internal customer services queries and provide author FAQs. 5. Capture data citation in reference list at point of submission in a structured way. 6. Data availability should be captured in a structured way. 7. Update XML DTD for data citation tagging. 8. Display data citations in the article. 9. Deliver data citation metadata to Cross-Ref.

How Nature Research journals guide the authors....

- Data availability statement - about the availability of the minimal dataset that would be necessary to interpret, replicate and build upon the methods or findings reported in the article.
- The preferred way to share large data sets is via public repositories.
- Unstructured repositories like figshare and Dryad are suitable alternatives if no structured public repositories exist.



Research data

Submitting to Research Data
Support

Who can use Research Data
Support

Pricing for Research Data
Support

Recommended Repositories

In general, data should be submitted to discipline-specific, community-recognised repository where possible, or to [generalist repositories](#) if no suitable community resource is available.

This list is derived from *Scientific Data's* recommended repository list. The list is also [available in figshare](#) for free reuse, with attribution, by others. *Scientific Data* is an open access data journal published by Springer Nature.

Research data

Find the best repository for your data

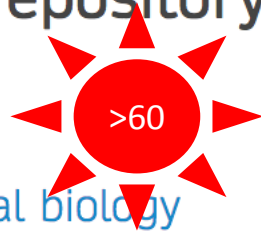
Biological sciences

Chemistry & chemical biology

Earth & environmental sciences

Generalist repositories

Health sciences



Materials science

Physics, astrophysics & astronomy

Social science

Other repositories

Raising awareness: Journal data policies

Policy Types



- Springer Nature launched a data policy standardisation initiative in 2016¹
- More than 1,400 (~60%) Springer Nature journals have adopted a standard research data policy
- Approach enables all journals to adopt a policy even if they are new to data sharing
- All policies and journals promote data citation in information for authors
- Similar initiatives since introduced at Elsevier, Wiley, Taylor & Francis
- Global collaboration on journal data policies via Research Data Alliance - a community-driven organization set up with the goal of building the social and technical infrastructure to enable open sharing of data.

1. *Standardising and harmonising research data policy in scholarly publishing*

Iain Hrynaszkiewicz, Aliaksandr Birukou, Mathias Astell, Sowmya Swaminathan, Amye Kenall, Varsha Khodiyar
International Journal of Digital Curation; doi: <https://doi.org/10.2218/ijdc.v12i1.531>

Data provides the core of your publication and is the foundation on which scientific, technical and medical knowledge is built. As a researcher, you are encouraged – and sometimes mandated – to store your research data and make it accessible and discoverable, so others can reuse it.

Elsevier journals have different guidelines when it comes to research data, either encouraging or requiring you take certain actions. Although much research data is disseminated as part of journal articles, a host of other data is not made available through article publication. These guidelines concern research data that often underlies, but exists outside of research articles. You can use the information below to understand the requirements of your chosen journal, so you can make informed choices about where and how to publish.

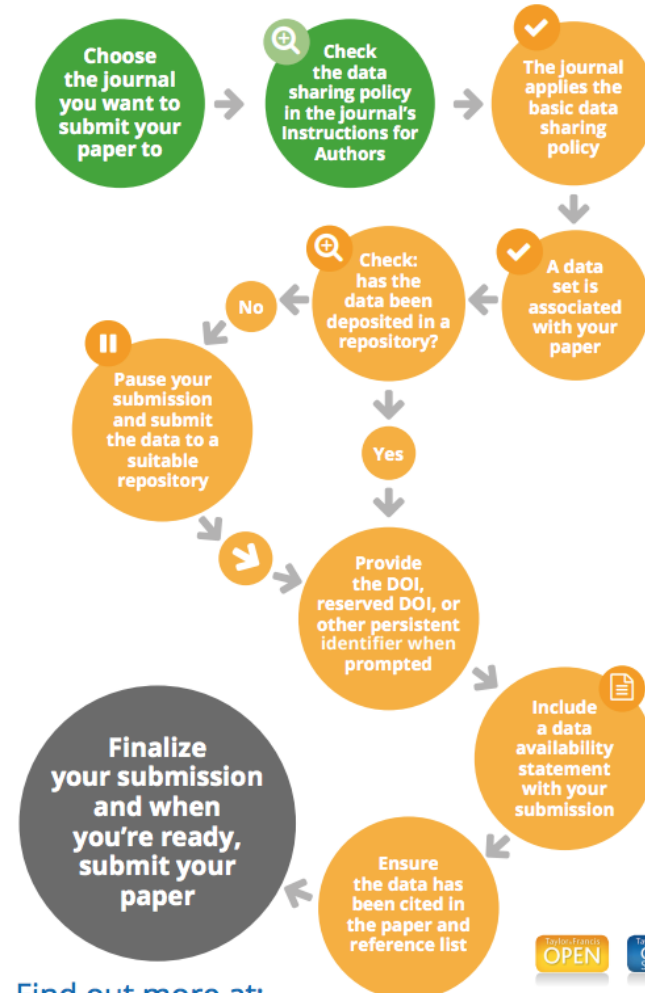


Sharing data: understanding our basic data sharing policy

- ✓ Submitting your article to a Taylor & Francis, Routledge or Cogent OA journal
- ✓ A data set is associated with your paper



The example below shows the steps you'll need to follow for many of our journals with a **basic data sharing policy**, which encourages authors to deposit data in a suitable repository, to cite it, and provide a data availability statement.



Data Sharing Policy

See how Wiley encourages authors to share their research data

Data Citation Policy

Check how research data should be cited

Data Sharing Service

Discover how your data could be automatically archived

Providing support: Advice and guidance

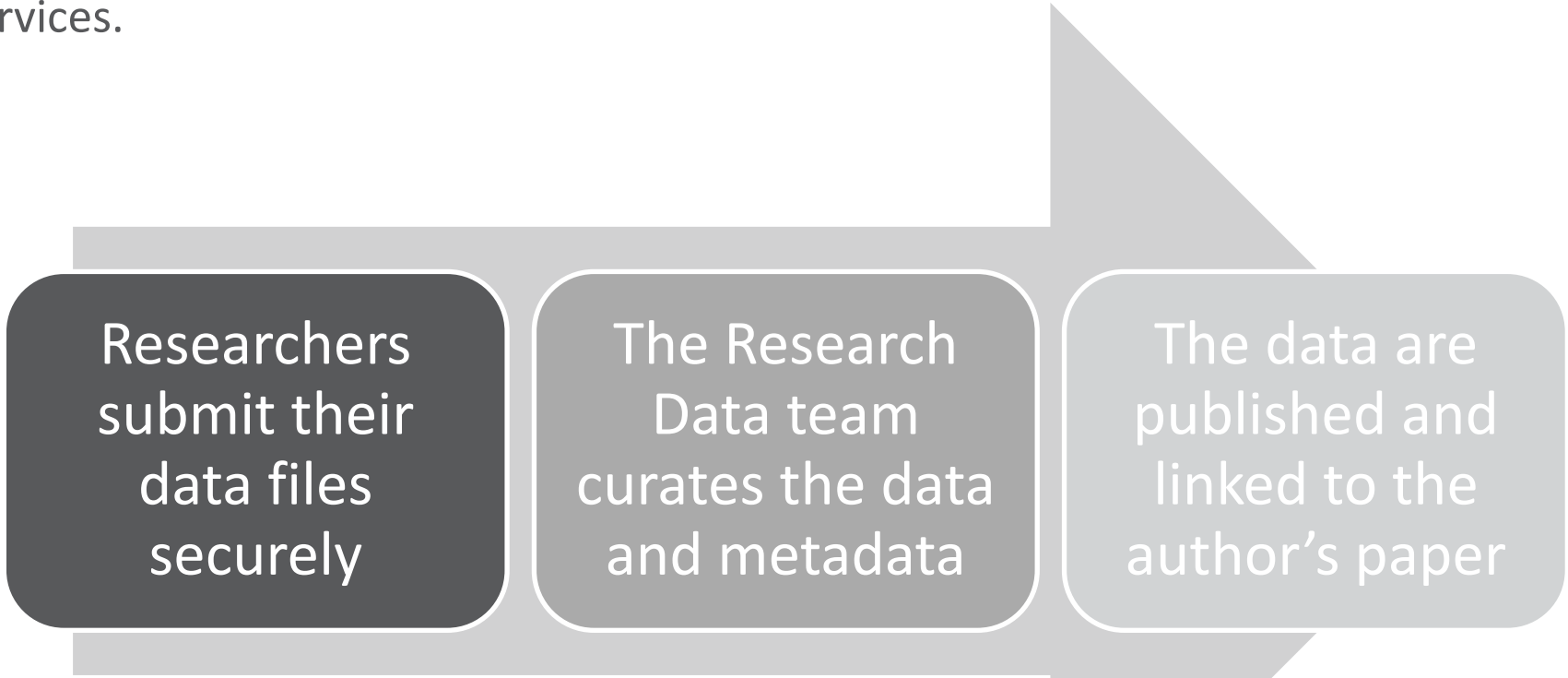
- **Research Data Helpdesk** researchdata@springernature.com
 - Run by members of the Springer Nature Data Publishing team
 - Expertise in data curation and management, archiving and digital preservation, copyright and licensing, Open Access publishing
- **Recommended repositories list**
 - More than 90 repositories; transparent criteria for approval
 - Supports community/discipline repositories
- **Guidance on writing data availability statements**
 - Contextual examples from all disciplines



Slide c/o Iain Hrynaszkiewicz

Hands-on assistance: Research Data Support

To help Springer Nature authors and journals follow good practice in sharing and archiving of research data, we offer optional data deposition and curation services.



Researchers
submit their
data files
securely

The Research
Data team
curates the data
and metadata

The data are
published and
linked to the
author's paper

More information is available on our website
<https://go.nature.com/ResearchDataServices>

Slide c/o Iain Hrynaszkievicz

natureresearch

ORGANISE AND SHARE YOUR DATA FASTER WITH RESEARCH DATA SUPPORT

Helping you make your data more findable,
accessible, understandable and citeable

What you need to know before submitting your data

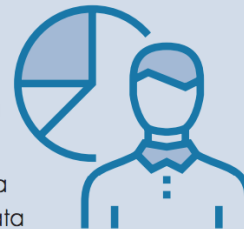
Can you share your data?

You need to know if you have permission to openly share the data. If you are acting on behalf of the data owner, make sure you have their agreement before submitting.



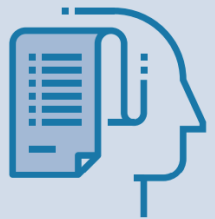
Are your data ready to share?

You need to have your data files ready for uploading, along with a short description of your data. Data cannot contain any identifiable human or sensitive data.



Who is the owner of the data?

When using this service the data owner (whether that is you, your institute, or someone else) retains full ownership and copyright for all files submitted.



Submission

How Research Data Support Editors will improve your data

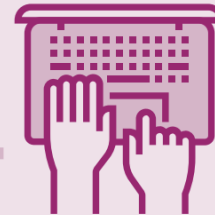
Generate persistent, citeable identifiers for datasets

A DOI (Digital Object Identifier) is generated for each dataset, enhancing accessibility and providing a persistent link for citation.



Increase findability by improving metadata and keywords

Metadata are checked and copy-edited so that they are clear, readable, and free from spelling and grammatical errors. Relevant keywords are added to enhance searchability.



Create clear, understandable descriptions of datasets

A comprehensive description of datasets and the underlying research method is created based on metadata, any associated manuscript(s) and input from the data owner.



Add relevant funder information

To ensure compliance with any funder requirements, funder information and acknowledgments are added to datasets information and metadata.



Link datasets to related research

Whenever possible, datasets are linked to associated publication(s) - allowing research that built on datasets to be identified and accessed.



★ The benefits to you and your data ★

An enhanced and fully citeable dataset

Upon completion, data can be made available at any time - for example in conjunction with a publication.



A complete metadata record

Completely machine- and human-readable record, which enables greater understandability, findability and reusability of your data.



Enhancement report

Detailed report outlining the enhancements made to your data and confirming it has been through the service.



Data availability statement

A complete data availability statement for your dataset will also be provided. This can be used to inform others of where and how to access your data, including in articles, books and other publications.





Research Data Support pilot with Wellcome Trust

In July 2018, the Wellcome Trust partnered with Springer Nature on a pilot to make our Research Data Support service available to all Wellcome-funded researchers at no cost to the researcher.

- Removes the barrier of lack of funding for research data sharing
- Researchers remain in control of what they publish and choose licence for their data
- Available for research data supporting any publication in any journal or publisher

Slide c/o Iain Hrynaskiewicz

Wellcome Open Research [SUBMIT YOUR RESEARCH](#)

BROWSE GATEWAYS HOW TO PUBLISH ABOUT BLOG

RESEARCH ARTICLE
Carbapenems as water soluble organocatalysts
 1; referees: awaiting peer review]

Thomas L. Williams, Alexander R. Nödling, Yu-Hsuan Tsai, Louis Y. P. Luk

[Author details](#)

Data availability

The H NMR spectra, chiral HPLC chromatograms and full protein LC-MS chromatograms, and mass spectra datasets generated during and/or analyzed during the current study are available in the figshare repository: <https://doi.org/10.6084/m9.figshare.6973880>

NMR spectra, chiral HPLC chromatograms and full protein LC-MS chromatograms and mass spectra related to carbapenems as water soluble organocatalysts

43 views | 0 downloads | 1 citations

READ THE PEER-REVIEWED PUBLICATION
 Carbapenems as water soluble organocatalysts

SPRINGER NATURE | Research Data Support

CATEGORIES

- Catalysis and Mechanisms of Reactions
- Biocatalysis and Enzyme Technology
- Analytical Biochemistry
- Biochemistry

KEYWORD(S)

- Organocatalysts
- water-soluble iridium catalyst
- carbapenem
- Michael addition
- artificial enzyme
- NMR spectroscopic analyses
- HPLC analysis
- high performance liquid chromatography-mass s...
- carbapenem antibiotics
- β -lactamase
- NMR
- ^1H NMR

LICENCE

EXPORT

- RefWorks
- BiTeX
- Ref. manager
- Mendeley
- Endnote
- DataCite
- NLM
- DO

FUNDING

We are grateful for support from the Cardiff School of Chemistry, Leverhulme Trust (RPG-2017-155 to LYPL) and Wellcome Trust (202056 to L.Y.P.L. and 200730 to Y.H.T.). We also thank EPSRC (EP/L027240/1) for acquisition of Waters Synapt G2-BI LC-MS system.

<http://dx.doi.org/10.12688/wellcomeopenres.14721.1>

DOI and suggested data citations provided

Useful and meaningful title

The data can be viewed in the browser without being downloaded

Altmetrics provides information on downloads and citations

Link to the peer reviewed article connected to this dataset

Relevant categories and keywords allow other researchers to find the dataset

Licence added to make reuse conditions clear

Comprehensive description including the data context of the study and data gathering method

Funder information available

Explore more content

Cite

Download (44.62 kB)

Share

Embed

+ Collect

Capture-mark-recapture data modelling survival rates of *Microcebus murinus* in relation to glucocorticoid level, parasite infection and body condition

01.09.2017, 09:26 by Josué H Rakotonirainy, Peter M Kappeler, Eva Kaesler, Anni M Härmäläinen, Clemens Kirschbaum, Cornelia Kraus

This dataset consists of an Excel spreadsheet containing capture-mark-recapture data, which were used to model survival of *Microcebus murinus* in different contexts.

These were:

- a Multistate modelling approach to model semi-annual survival relative to hair cortisol concentration (HCC) and scaled mass index (SMI). Median or third quartile were used as categorization cut-off.
- a Cormack-Jolly-Seber (CJS) modelling approach to model survival over the reproductive season relative to hair cortisol concentration, scaled mass index, and pattern of parasitism which was measured as the parasite species richness (number of distinct parasite morphotypes found per individual), the multiple infection (presence of more than one parasite morphotype), and the overall parasite prevalence (presence of at least one parasite morphotype).

The data used to assess the link between semi-annual survival rates and HCC includes the results of capture sessions held in October 2012, 2013, 2014, April 2013, and March 2014, during which a total of 171 individuals (74 females, 97 males) were captured. The same dataset, excluding the October 2014 session, was used to assess the effect of SMI on semi-annual survival probabilities, for a total of 149 individuals (83 females, 86 males). The dataset used for the CJS models includes data collected during monthly trapping sessions between September 2012 and April 2013, for a total of 48 individuals (16 females, 32 males).

All research activities conducted in Madagascar received official approval from the Ministère de l'Environnement, de l'Écologie, and de la Mer et des Forêts, and comply with national animal care legislation of Madagascar.

FUNDING

Deutscher Akademischer Austauschdienst (A/12/90426) and Deutsche Forschungsgemeinschaft (KR 3634/4-1)

RESEARCH DATA SUPPORT

Research data support provided by Springer Nature.

COMMENTS (0)

Write your comment here ...

Post comment

840 views

18 downloads

1 citations



READ THE PEER-REVIEWED PUBLICATION:

Hair cortisol concentrations correlate negatively with survival in a wild primate population

SPRINGER NATURE

CATEGORIES

- Parasitology
- Behavioural Ecology
- Ecological Physiology
- Physiology
- Population Ecology

KEYWORD(S)

- Microcebus murinus
- Cormack-Jolly-Seber models
- Survival Analysis
- Glucocorticoid Function
- Capture mark-recapture
- Mouse Lemur Primates
- Hair cortisol levels
- stress agents
- parasitisms
- ecological stressors
- Madagascar
- Cort-fitness hypothesis
- Multistate Survival Models
- Multistate models

LICENCE

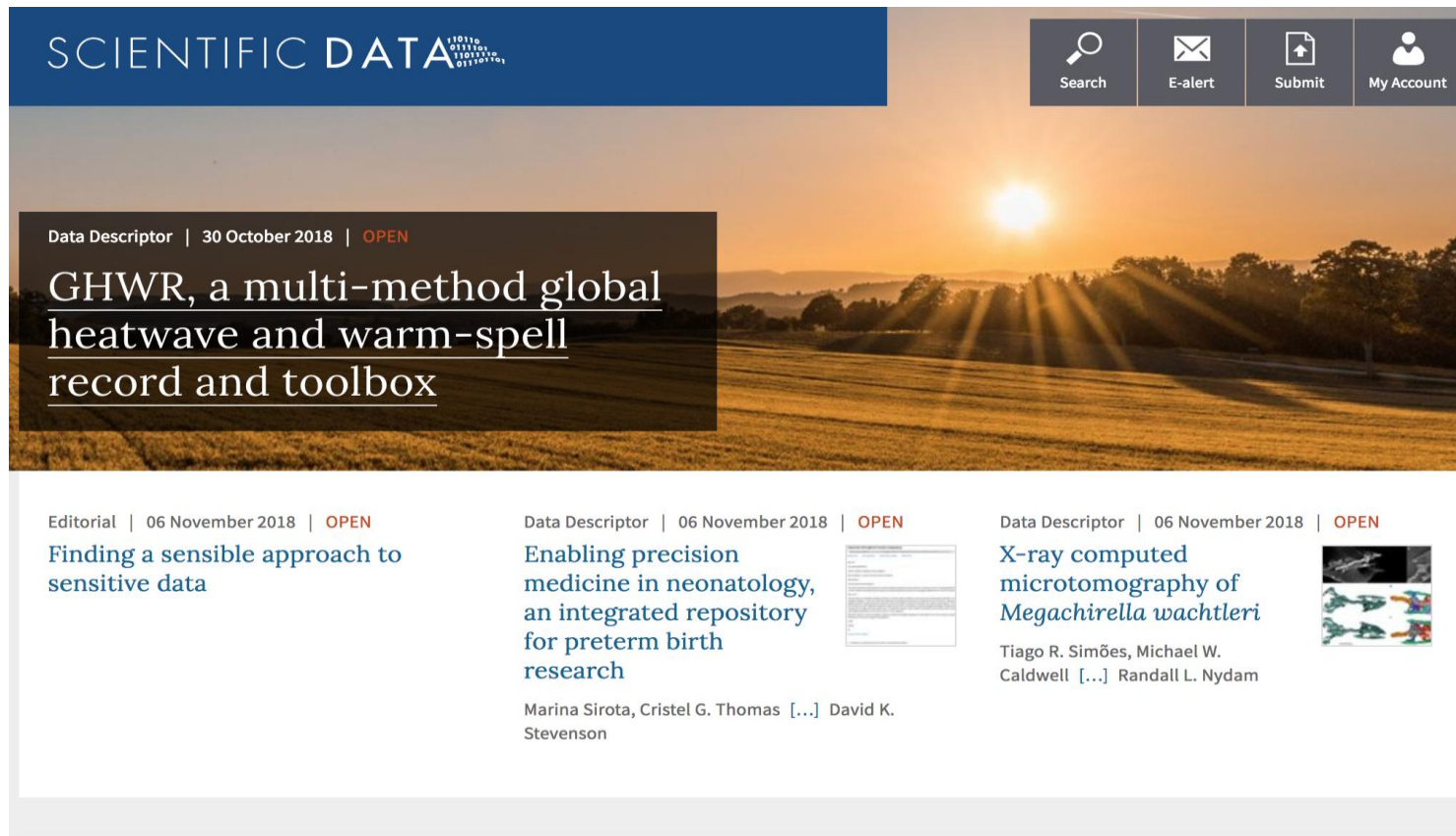


EXPORT

RefWorks

BibTeX

From the Nature Research portfolio: *Scientific Data*



Scientific Data is an open-access, online-only journal for descriptions of scientifically valuable datasets.

The articles, known as Data Descriptors, combine traditional narrative content with curated, structured descriptions (metadata) of the published data to provide a new framework for data-sharing and –reuse.

The Data Descriptor article

Peer-reviewed articles enabling authors to provide comprehensive methodological detail about their data. Does not contain tests of new scientific hypotheses.

Sections:

- Title
- Abstract
- Background & Summary
- **Methods**
- **Technical Validation**
- **Data Records** ←
- **Usage Notes**
- Figures & Tables
- References
- **Data Citations** ←

Data Records

All the samples used in this study are summarized in Table 1. Consistent identifiers are used in Tables 2 and 3 to allow mapping between the proteomic and transcriptomic data outputs.

Data Record 1

The raw data, peaklists (.mgf), ProteomeDiscoverer result files (.msf) and ProteomeDiscoverer workflow files (.xml) have been uploaded to ProteomeXchange (<http://www.proteomexchange.org/>) with the following accession number PXD000134 (ref. 67; Table 2).

Data Record 2

Microarray data are available at the NCBI Gene Expression Omnibus (GEO) database under the accession numbers GSE26451 (ref. 68) and GSE26453 (ref. 69; Table 3).

Data Record 3

The peptide and protein identification data sets have been annotated by The Global Proteome Machine at <http://gpsdb.thegpm.org/>

Data Record 4

The peptide and protein identification data sets have been annotated by the StemCellOmicsRepository (SCOR) at <http://scor.chem.wisc.edu/>

Data Citations

67. Low, T. Y. *et al.* ProteomeXchange: PXD000134 (2013).

68. Chin, A. *et al.* Gene Expression Omnibus: GSE26451 (2011).

69. Chin, A. *et al.* Gene Expression Omnibus: GSE26453 (2011).



Search



E-alert



Submit



My Account

Policies

[Editorial & Publishing Policies](#)[For Referees](#)[Data Policies](#)[Recommended Data Repositories](#)

Recommended Data Repositories

Scientific Data mandates the release of datasets accompanying our Data Descriptors, but we do not ourselves host data. Instead, we ask authors to submit datasets to an appropriate public data repository. Data should be submitted to discipline-specific, community-recognized repositories where possible, or to [generalist repositories](#) if no suitable community resource is available.

Repositories included on this page have been evaluated to ensure that they meet our requirements for data access, preservation and stability. Please be aware, however, that some repositories on this page may only accept data from those funded by specific sources, or may charge for hosting data. Please ensure you are aware of any deposition policies for your chosen repository. If your repository of choice is not listed please see our [guidelines for suggesting additional repositories](#).

<http://www.copdess.org/enabling-fair-data-project/>

Commitment to Enabling FAIR Data in the Earth, Space, and Environmental Sciences

[READ MORE](#)

Publication of scholarly articles in the Earth, space, and environmental science community is conditional upon the concurrent availability of the data underpinning the research finding, with only a few, standard, widely adopted exceptions, such as around privacy for human subjects or to protect heritage field samples. These data should, to the greatest extent possible, be shared, open, and stored in community-approved FAIR-aligned repositories. Leading repositories provide additional quality checks around domain data and data services and facilitate discovery and reuse of data and other research outputs.

This effort builds on the work of The Coalition on Publishing Data in the Earth and Space Sciences (COPDESS.org), Earth Science Information Partners (ESIP), Research Data Alliance (RDA), the scientific journals, and domain repositories to ensure that well documented data, preserved in a repository with community agreed-upon metadata, and supporting persistent identifiers becomes part of the expected research products submitted in support of each publication. **This is a community-driven project that will be managed by both ESIP and RDA after the grant period concludes.**

<http://www.copdess.org/enabling-fair-data-project/>

Science Data Communities

- The American Geophysical Union (AGU)
- European Geosciences Union (EGU)
- Earth Science Information Partners (ESIP)
- Research Data Alliance (RDA)
- EarthCube / Council for Data Facilities
- Coalition on Publishing Data in Earth and Space Sciences (COPDESS.org)
- FORCE11

Publishers

- American Geophysical Union (AGU)
- Proceedings of the National Academy of Sciences (PNAS)
- Nature
- Science/AAAS
- Elsevier
- PLOS
- Hindawi
- Copernicus Publications
- Wiley

Repositories and Data Ecosystem Infrastructure

- National Computational Infrastructure (NCI)
- AuScope
- Australian Research Data Commons (ARDC)
- Center for Open Science
- DataCite/Re3Data
- ORCID
- CrossRef
- CHORUS
- Scholix
- OSGeo
- Pangaea
- DataONE
- World Data System

...AND GROWING

Beyond data....

Aiding transparency of research reporting at Nature Research

- Raise standards of **reporting on** statistics
- Ensure reporting on key elements of **experimental design**
- Information on **materials** used
- Eliminate length restrictions for **Methods** sections
- Encourage provision of **source data**
- **Independent analysis in 2017 finds the checklist is resulting in improved reporting of risks of bias, randomisation, blinding, exclusions and sample size calculations**

<http://dx.doi.org/10.1101/187245>

Slide c/o Iain Hrynaszkiewicz

the information can be found.

	Reported in section/paragraph or page #:
<p>► Animal Models</p> <p>8. Report species, strain, sex and age of animals</p> <p>9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.</p> <p>10. We recommend consulting the ARRIVE guidelines (PLoS Biol. 8(6), e1000412,2010) to ensure that other relevant aspects of animal studies are adequately reported.</p>	<p></p> <p></p> <p></p>

Nature Research editorial policies for better science

- Before peer review, the corresponding author must complete an [editorial policy checklist](#) to ensure compliance with Nature Research editorial policies.
- Authors of research articles in the [life sciences](#), behavioural & social sciences and ecology, evolution & environmental sciences are required to provide details about elements of experimental and analytical design that are frequently poorly reported in a [reporting summary](#) that will be made available to editors and reviewers during manuscript assessment. The reporting summary will be published with all accepted manuscripts.
- For [physical sciences](#), we require authors of research articles in some specific areas to provide details of characterization, or experimental and analytical design in a reporting summary, which will be made available to editors and reviewers during manuscript assessment and published with an accepted manuscript: Solar cells and Claims of lasing
- Details of [geological samples and palaeontological specimens](#) should include clear provenance information to ensure full transparency of the research methods. Any submission detailing new material from protected sites should include information regarding the requisite permission obtained. Palaeontological and type specimens should be deposited in a recognised museum or collection to permit free access by other researchers in perpetuity.

Materials and protocols

- A condition of publication in a Nature Research journal is that authors are required to make unique materials promptly available to others without undue qualifications.
- Nature Research supports the [Resource Identification Initiative](#), with the aim of promoting unique, persistent identification and tracking of key biological resources, including antibodies, cell lines, model organisms and tools. We encourage authors to include unique identifiers provided by the [Resource Identification Portal](#), (RRIDs; for example, Antibody: [RRID:AB 2140114](#); Organism: [RRID:MGI MGI:3840442](#)), in the manuscript.

Code availability

- A subset of Nature journals undertake peer review of custom code and software when it is central to the manuscript. In those cases, editors will request authors must fill out a [Code and Software submission checklist](#).
- Upon publication, Nature Research Journals consider it best practice to release custom computer code in a way that allows readers to repeat the published results. We also strongly encourage authors to deposit and manage their source code revisions in an established software version control repository, such as [GitHub](#), and to use a DOI-minting repository to uniquely identify the code and provide details of its license of use. We strongly encourage the use of a [license](#) approved by the open source initiative.

Partnership with Code Ocean - a computational reproducibility platform

A Code Ocean compute capsule allows an author to configure their computational environment precisely, and guarantees reviewers access to that **exact same environment** — without needing to download or install anything;

Code Ocean is designed to run multiple programming languages in one environment with the click of a single button;

Publishers and researchers can link a published capsule as a “version of record,” with an associated DOI, to an associated article, thereby enabling readers to reproduce published findings with the click of a button;

Perhaps most importantly, Code Ocean guides users towards creating one canonical set of ‘Published Results’, and making transparent from start to finish how those results were generated.



Thank you

m.skipper@nature.com

@magda_skipper

Special thanks go to:

- Iain Hrynaszkiewicz, Head of Data Publishing
- Andrew Hufton, Chief Editor, *Scientific Data*
- Sowmya Swaminathan, Head of Editorial Policy and Research Integrity

The story behind the image



Antarctica meltdown could double sea level rise

Researchers at Pennsylvania State University have been considering how quickly a glacial ice melt in Antarctica would raise sea levels. By updating models with new discoveries and comparing them with past sea-level rise events they predict that a melting Antarctica could raise oceans by more than 3 feet by the end of the century if greenhouse gas emissions continued unabated, roughly doubling previous total sea-level rise estimates. Rising seas could put many of the world's coastlines underwater or at risk of flooding and storm surges.