World Data System
Webinar Series 2024

Joint Statement on Research Data

20 June 2024
13:00 UTC/9:00 AM EST

DOI: 10.5281/zenodo.12570909

WORLD DATA SYSTEM

Explore the Path to Proper Participation:
Visit Our Etiquette Page!

- Go to worlddatasystem.org/virtual-meeting-etiquette or scan the QR code

- We welcome your questions in the Q&A section

- Please note that this webinar will be recorded

WDS Virtual Meeting Etiquette

Speakers

Helena Cousijn

DataCite
CONNECTING RESEARCH, ADVANCING KNOWLEDGE

Martyn Rittman

Crossref

Hylke Koers

STM

Reyna Jenkyns

ITO

Agenda

5 min	Welcome and speaker introductions (Meredith Goins, WDS-IPO)
25 min	Hylke Koers, STM - Introduction on what is the Joint Statement on Research Data and History of Joint Statement on Research Data • Crossref perspective by Martyn Rittman • DataCite perspective by Helena Cousijn Examples from stakeholder types (researchers, data repositories, publishers, funders, research organizations, policymaking institutions)
10 min	Why WDS signed on (Reyna Jenkyns, WDS-ITO)
5 min	Call to action to attendees to sign on
15 min	Q&A (Meredith Goins, WDS-IPO)

140+ members

Academic publishers focused on science, technology, medicine, social sciences, and humanities

66% of articles

Our members collectively publish 66% of all journal articles

20 countries

STM spans the globe — made up of publishers of all shapes and sizes

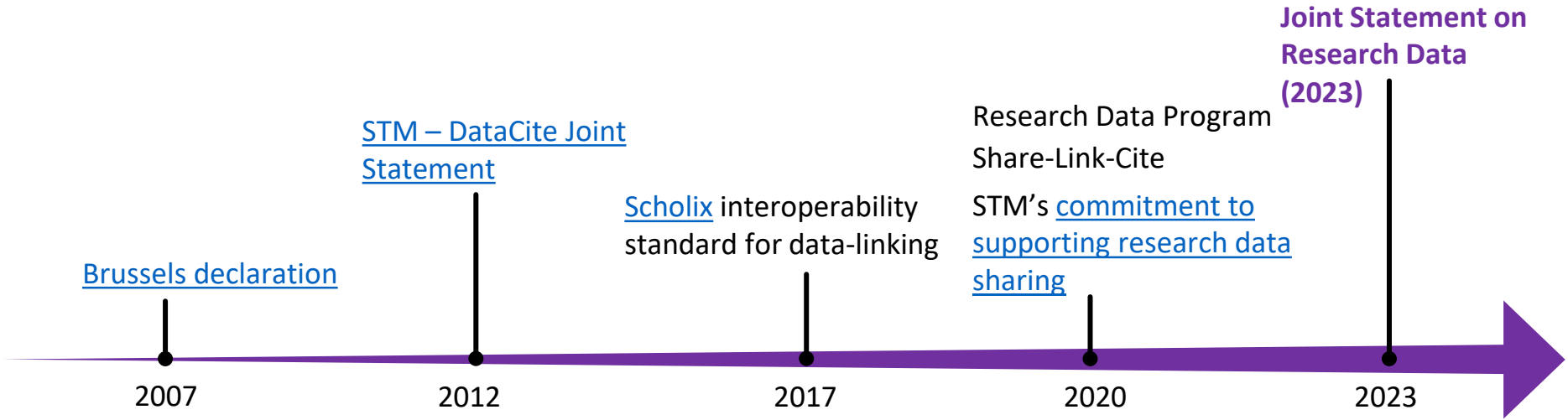
About STM

STM is the standard bearer for the academic publishing industry, working with our members to advance trusted research worldwide. We are committed to ensuring that the great discoveries of our time are communicated with pinpoint accuracy, clarity and integrity. We champion innovation across academic research, stimulating the development of new technologies and guidance on universal standards.

STM Advancing trusted research

STM has a long-standing commitment to fostering the sharing of research data

Sharing research data is crucial to advance science and research. It advances transparency, reproducibility, and provides additional opportunities for scientific discovery and collaboration. STM is committed to supporting data sharing to advance open scholarship.



STM

STM RESEARCH DATA PROGRAM

21 PARTICIPATING PUBLISHERS

13,064 PARTICIPATING JOURNALS

ELSEVIER · WILEY · SPRINGER NATURE · IOP Publishing · OSA The Optical Society 100 Since 1916 · JAMA Network · emerald PUBLISHING · CAMBRIDGE UNIVERSITY PRESS · DE GRUYTER · Taylor & Francis Group · AMERICAN PSYCHOLOGICAL ASSOCIATION · AGU ADVANCING EARTH AND SPACE SCIENCE · ACS Chemistry for Life® · SAGE Publishing · BRILL · ASCO AMERICAN SOCIETY OF CLINICAL ONCOLOGY KNOWLEDGE CONQUERS CANCER · OXFORD UNIVERSITY PRESS · Science JOURNALS · Karger · Wolters Kluwer

JOURNALS: AVERAGE NUMBER OF DATA POLICIES

29% START OF 2020

52% END OF 2020

OUR PARTNERS

COS CENTER FOR OPEN SCIENCE · FAIRSFAIR · Crossref · ripeta · CHORUS · RDA RESEARCH DATA ALLIANCE

15 MEETINGS, WEBINARS & WORKSHOPS

NIST National Institute of Standards and Technology U.S. Department of Commerce · European Commission · TU Delft · NIH National Institutes of Health · CESAER · EMSL · DANS · VT VIRGINIA TECH · TIB

ARTICLES THAT CONTAIN A DATA AVAILABILITY STATEMENT

7% START OF 2020

15% END OF 2020

JOURNALS THAT CONNECT TO SCHOLIX

34% Q2 OF 2020

43% END OF 2020

https://www.stm-assoc.org/research-data-program/

The publishing partner in shaping the research communication system of the future

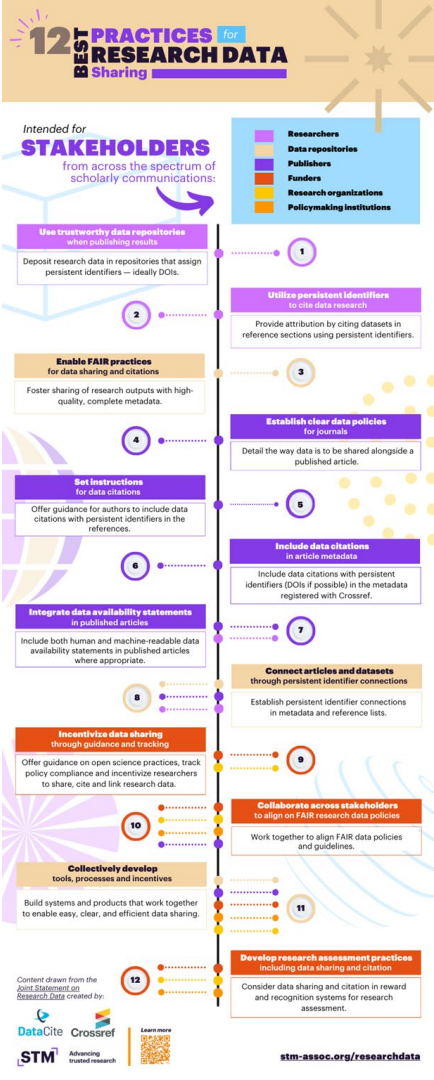
STM

The 2023 Joint Statement on Research Data

- Re-affirms the importance of research data to improve the utility and rigor of the scholarly record
- Significant progress since the original 2012 statement:
  - 10 million data citations tracked;
  - thousands of repositories adopting data citation best practices;
  - thousands of journals adopting data policies, data availability statements and establishing persistent links between articles and datasets; and
  - the introduction of data policies by an increasing number of funders
- However, research data sharing is not yet a self-evident step in the research lifecycle.

We now have sufficient scholarly communication infrastructure in place to bring about widespread change and believe momentum is building for collective action

Aim of the statement: To accelerate adoption of best practices and policies, and encourage further development of critical policies in collaboration with a wide group of stakeholders



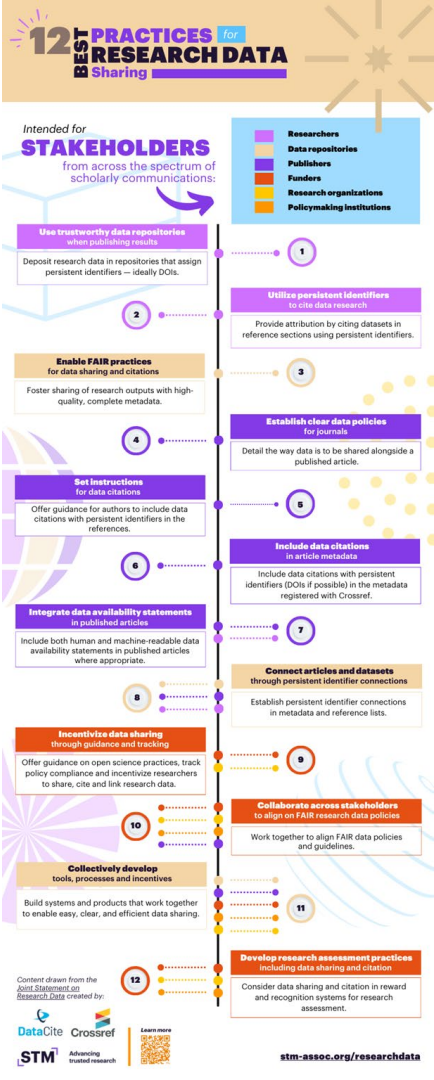
The 2023 Joint Statement on Research Data

Twelve best-practice recommendations directed at various stakeholder groups:

- Researchers
- Data repositories
- Publishers
- Funders
- Research organisations
- Policymaking institutions

With a call for individuals & organisations to endorse and explain how they will further these recommendations.

https://www.stm-assoc.org/research-data-program/bestpractices/



STM

Recommendations for publishers

- Publishers set appropriate **journal data policies**, describing the way in which data is to be shared alongside the published article.
- Publishers set **instructions for authors** to include **Data Citations** with persistent identifiers in the references section of articles.
- Publishers **include Data Citations and links to data in Data Availability Statements** with persistent identifiers (DOIs where available) in the **article metadata registered with Crossref**.
- In addition to Data Citations, **Data Availability Statements** (human- and machine-readable) are included in published articles where appropriate.
- Repositories and publishers **connect articles and datasets** through persistent identifier connections in the metadata and reference lists.

Recommendations for publishers

- Publishers set appropriate **journal data policies**, describing the way in which data is to be shared alongside the published article.
- Publishers set **instructions for authors** to include **Data Citations** with persistent identifiers in the references section of articles.
- Publishers **include Data Citations and links to data in Data Availability Statements** with persistent identifiers (DOIs where available) in the **article metadata registered with Crossref**.
- In addition to Data Citations, **Data Availability Statements** (human- and machine-readable) are included in published articles where appropriate.
- Repositories and publishers **connect articles and datasets** through persistent identifier connections in the metadata and reference lists.

STM

Recommendations for publishers

- Publishers set appropriate **journal data policies**, describing the way in which data is to be shared alongside the published article.
- Publishers set **instructions for authors** to include **Data Citations** with persistent identifiers in the references section of articles.
- Publishers **include Data Citations and links to data in Data Availability Statements** with persistent identifiers (DOIs where available) in the **article metadata registered with Crossref**.
- In addition to Data Citations, **Data Availability Statements** (human- and machine-readable) are included in published articles where appropriate.
- Repositories and publishers **connect articles and datasets** through persistent identifier connections in the metadata and reference lists.

Recommendations for publishers

- Publishers set appropriate **journal data policies**, describing the way in which data is to be shared alongside the published article.
- Publishers set **instructions for authors** to include **Data Citations** with persistent identifiers in the references section of articles.
- Publishers **include Data Citations and links to data in Data Availability Statements** with persistent identifiers (DOIs where available) in the **article metadata registered with Crossref**.
- In addition to Data Citations, **Data Availability Statements** (human- and machine-readable) are included in published articles where appropriate.
- Repositories and publishers **connect articles and datasets** through persistent identifier connections in the metadata and reference lists.

Recommendations for publishers

- Publishers set appropriate **journal data policies**, describing the way in which data is to be shared alongside the published article.
- Publishers set **instructions for authors** to include **Data Citations** with persistent identifiers in the references section of articles.
- Publishers **include Data Citations and links to data in Data Availability Statements** with persistent identifiers (DOIs where available) in the **article metadata registered with Crossref**.
- In addition to Data Citations, **Data Availability Statements** (human- and machine-readable) are included in published articles where appropriate.
- Repositories and publishers **connect articles and datasets** through persistent identifier connections in the metadata and reference lists.

STM

Recommendations for publishers

- Publishers set appropriate **journal data policies**, describing the way in which data is to be shared alongside the published article.
- Publishers set **instructions for authors** to include **Data Citations** with persistent identifiers in the references section of articles.
- Publishers **include Data Citations and links to data in Data Availability Statements** with persistent identifiers (DOIs where available) in the **article metadata registered with Crossref**.
- In addition to Data Citations, **Data Availability Statements** (human- and machine-readable) are included in published articles where appropriate.
- Reposito
persister

Progress will be monitored through STM Research Data Implementation Group

STM

Crossref, Data Citation, and the Research Nexus

Martyn Rittman, PhD
mrittman@crossref.org

Crossref

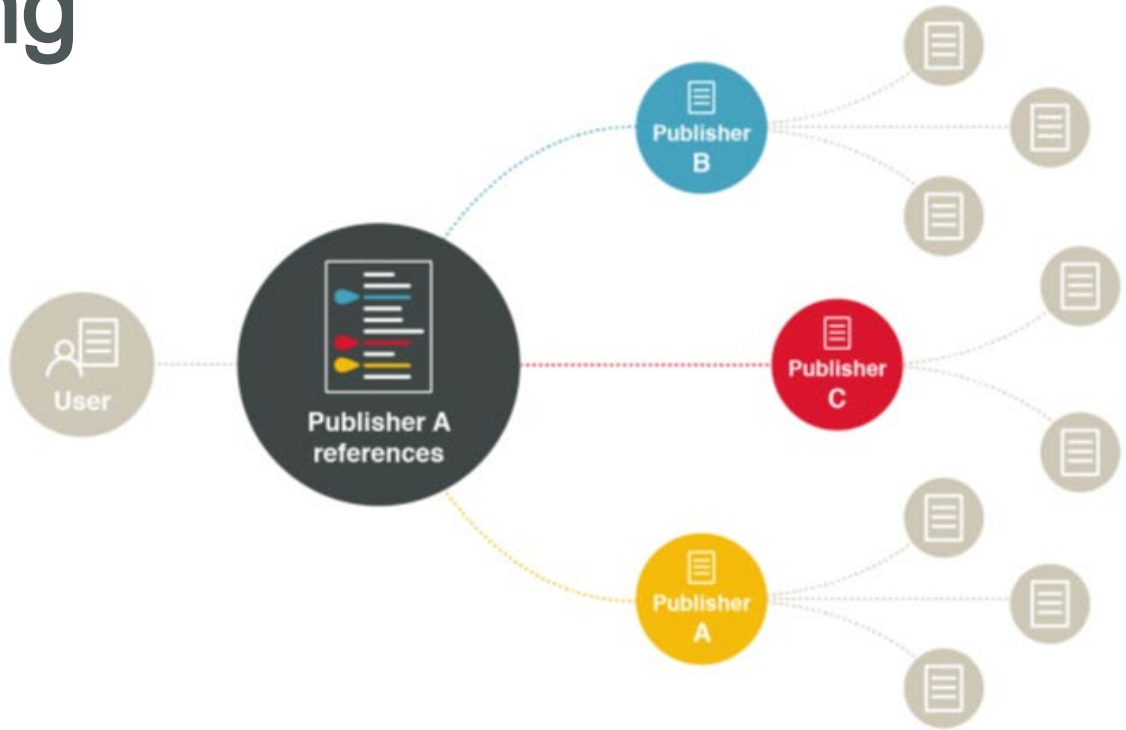
@CrossrefOrg

Our mission

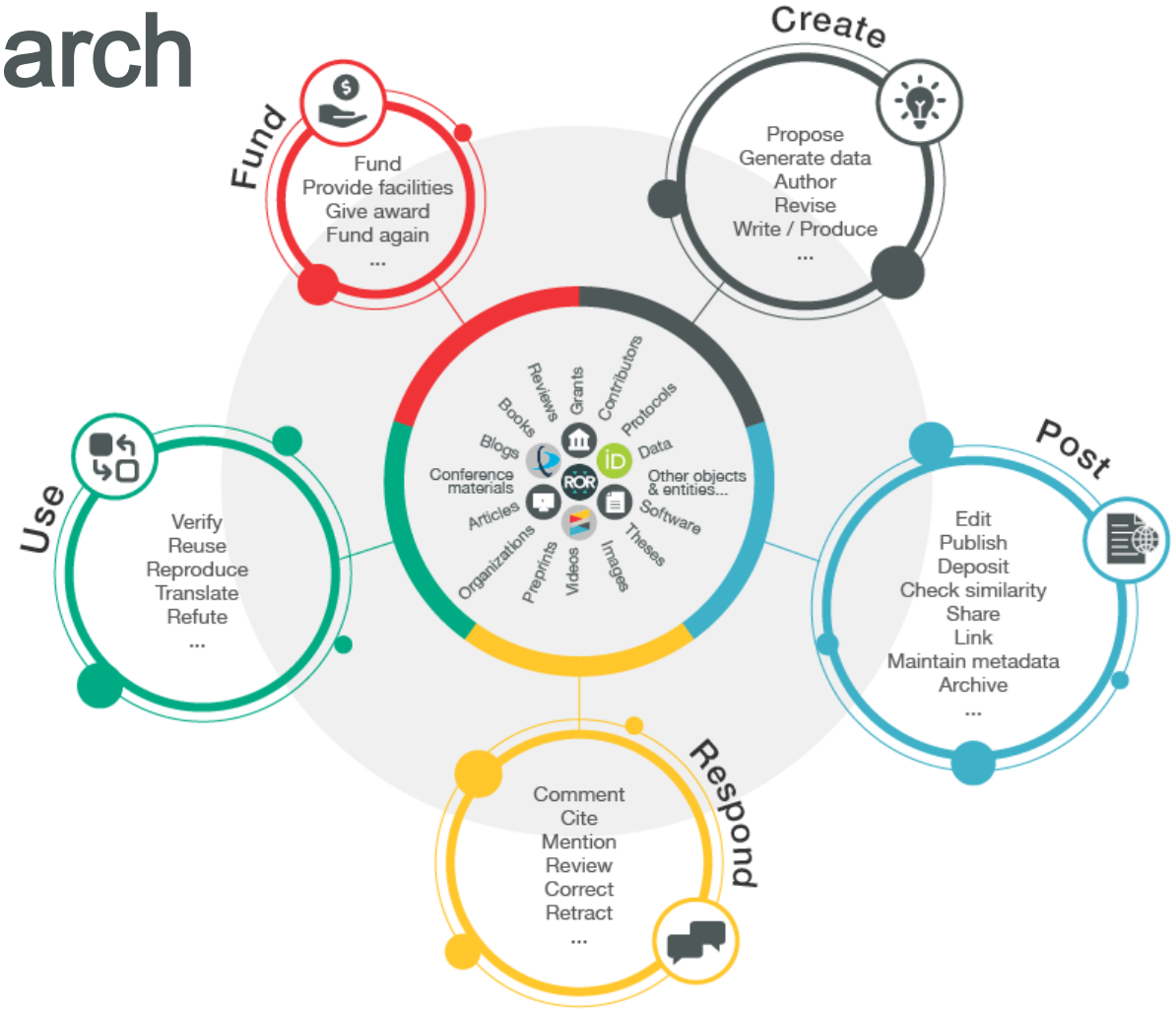
Crossref makes research outputs easy to find, cite, link, assess, and reuse.

We're a not-for-profit membership organization that exists to make scholarly communications better.

Reference Linking



The research nexus



Data citations

- 340,000 data citations in Scholix endpoint *
- >32 orgs with > 1000 data citations since 2023

*we know there are more than this

Data citations in reference lists

Include data citations in the references you register with Crossref. Can be as simple as just providing the DOI for the dataset:

```
<citation key="ref2">
  <doi>10.6084/m9.figshare.5981968</doi>
</citation>
```

Crossref recognizes that the DOI is for a DataCite DOI and pass it to Event Data.

Make sure your typesetters, copy editors and hosting platforms are briefed on how to work with these links and don’t remove them.

In practice

```
</citation>
▼<citation key="ref15">
  <doi>10.5281/zenodo.34014</doi>
</citation>
▼<citation key="ref16">
```

31. Martin, P. *et al*. Project IPAD, a database to catalogue the analysis of Fukushima Daiichi accident fragmental release material. *Mendeley* https://doi.org/10.17632/nz6hjbfs65.3 (2020).

```
378–381 (1997).</unstructured_citation>
</citation>
▼<citation key="626_CR31">
  <author>P Martin</author>
  <cYear>2020</cYear>
  <doi>10.17632/nz6hjbfs65.3</doi>
  <unstructured_citation>Martin, P. et al. Project IPAD, a database to catalogue the analysis of Fukushima Daiichi accident fragmental release material. Mendeley
  https://doi.org/10.17632/nz6hjbfs65.3 (2020).</unstructured_citation>
</citation>
```

But in some cases

11. SAPEA, Science Advice for Policy by European Academies (2019). *A Scientific Perspective on Microplastics in Nature and Society*. Berlin: SAPEA. doi: https://doi.org/10.26356/microplastics

```
<volume_title>A Scientific Perspective on Microplastics in Nature and Society</volume_title>
<author>SAPEA, Science Advice for Policy by European Academies</author>
<cYear>2019</cYear>
</citation>
```

8. Holinski A, Burke ML, Morgan SL, *et al.*: Biocuration - mapping resources and needs - Underlying Data (Version 2) [Data set].*Zenodo*. 2020. http://www.doi.org/10.5281/zenodo.3991737

```
</citation>
<citation key="ref-8">
  <journal_title>Zenodo.</journal_title>
  <author>A Holinski</author>
  <cYear>2020</cYear>
  <article_title>Biocuration - mapping resources and needs - Underlying Data (Version 2) [Data set].</article_titl
</citation>
```

Data citations as relationships

Include relationships between DOIs and other items:

```
<rel:related_item>
  <rel:description>Sicard-2018-External-database-S1
  </rel:description>
  <rel:inter_work_relation identifier-type="doi" relationship-
  type="isSupplementedby">
   10.6084/m9.figshare.5981968
  </rel:inter_work_relation>
</rel:related_item>
```

Relations metadata is currently available via Crossref's REST API and Event Data.

Delivering data citations

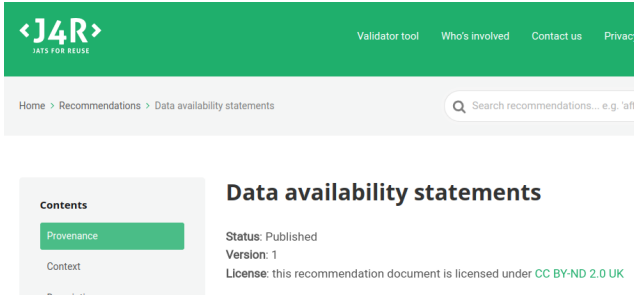
Scholix / Event Data:

- Not at production level
- Incomplete data
- Paused work on a successor in April 2024

Reference lists from REST API:

- Production support
- Complete and accurate
- Data citations are needles in a haystack (1 million in 1.5 billion)

Future schema changes



https://jats4r.niso.org/data-availability-statements/

Expanded citation markup with publication types:

```
<citation key="ref3"
publication_type="data">
```

Further plans to collect data availability statements in the Crossref metadata.

About DataCite

We are a global community that shares a common interest: to ensure that research outputs and resources are openly available and connected so that their reuse can advance knowledge across and between disciplines, now and in the future.

As a global non-profit membership organization, we work with 3100+ repositories in the world to provide DOIs for research outputs and resources.

Our community

DataCite

1406

Organizations

3100

Repositories

55

Countries

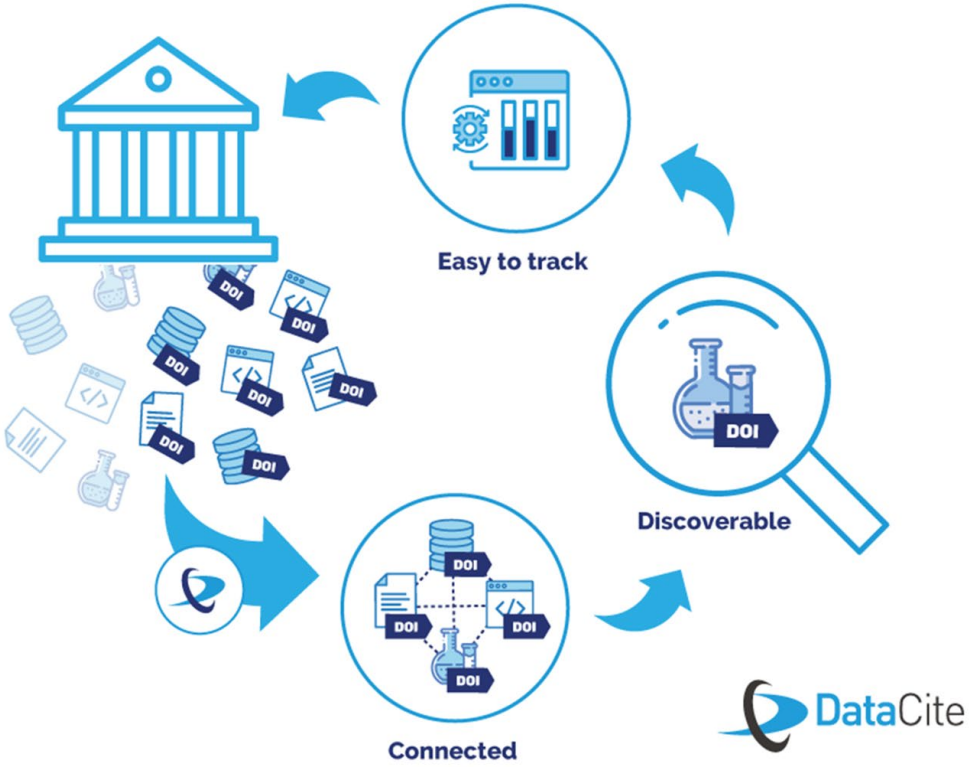
61m+

DOIs

(May 2024)

What we do & why

DataCite



Joint Statement on Research Data: a role for everyone

DataCite

Researchers

- When publishing their results, researchers deposit related research data and outputs in a trustworthy data repository that assigns persistent identifiers (DOIs where available). Researchers link to research data using persistent identifiers.
- When using research data created by others, researchers provide attribution by citing the datasets in the reference section using persistent identifiers.

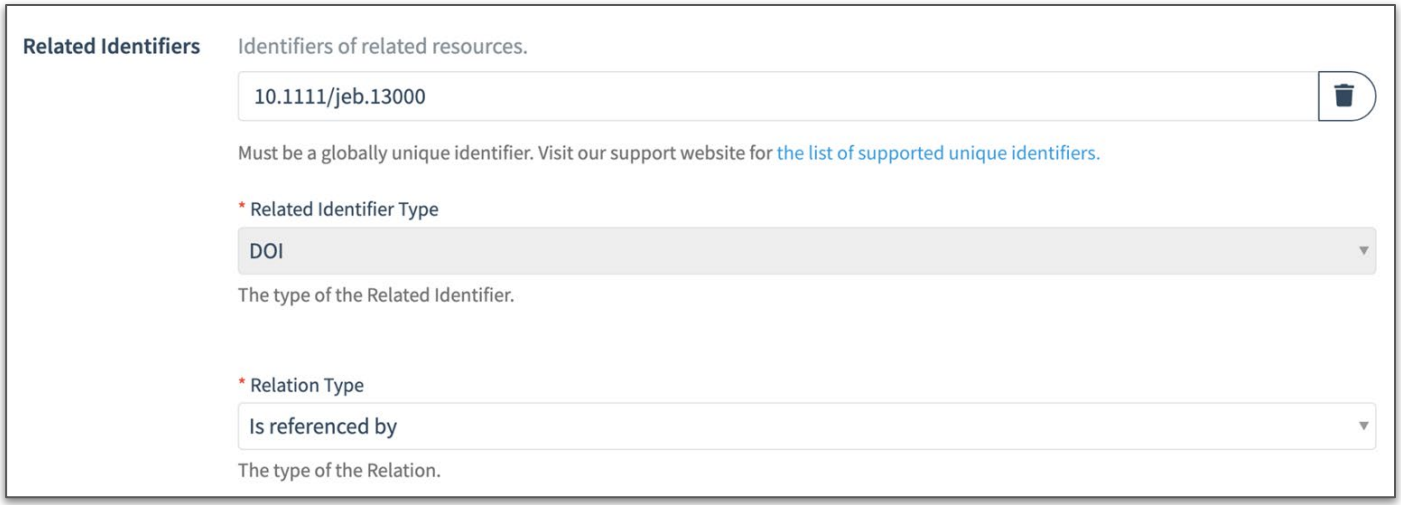
Repositories

- Data repositories enable sharing of research outputs in a FAIR way, including support for metadata quality and completeness.
- Repositories and publishers connect articles and datasets through persistent identifier connections in the metadata and reference lists.

Data Citation & Repositories

Repositories play an important role in data citation by linking data to related articles (or other outputs) and making this information available to the community

- Ask researchers to include associated publications in dataset metadata
- Include those links in DataCite metadata

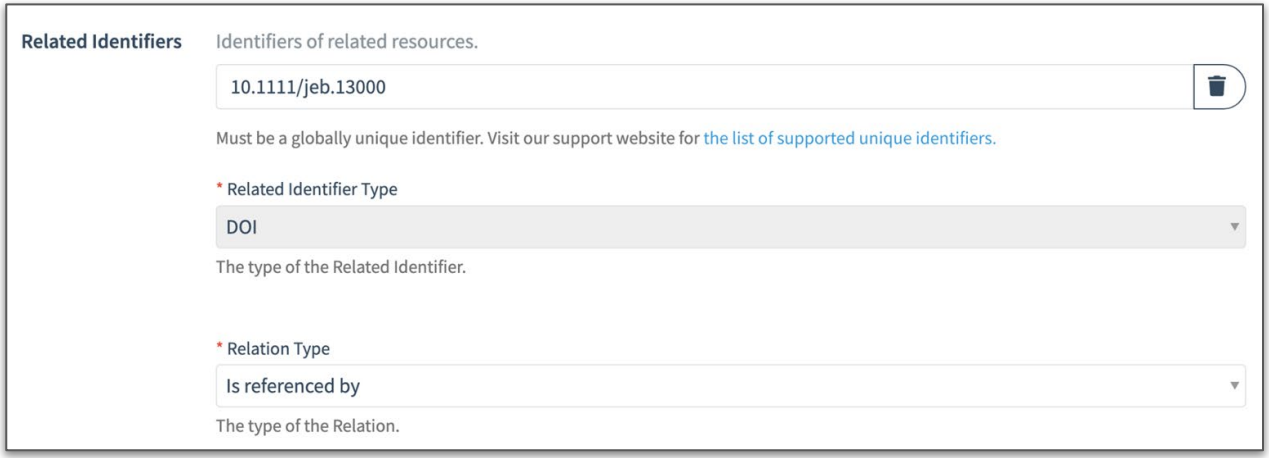


DataCite metadata deposit

The **RelatedIdentifier** field in the DataCite metadata schema can be populated in order to generate a link between the data and the publication that cites it.

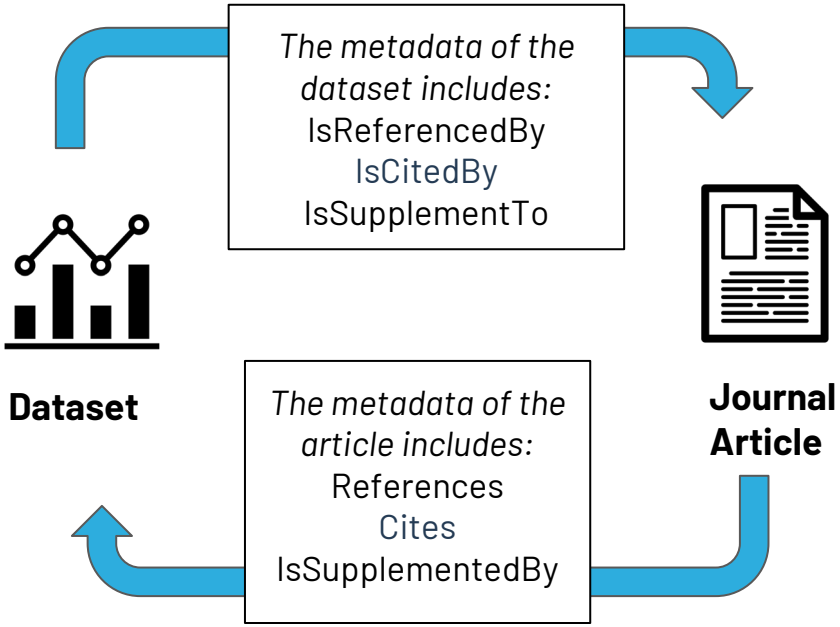
Mandatory sub-properties:

- **relatedIdentifierType**: the type of identifier
- **relationType**: the type of relationship

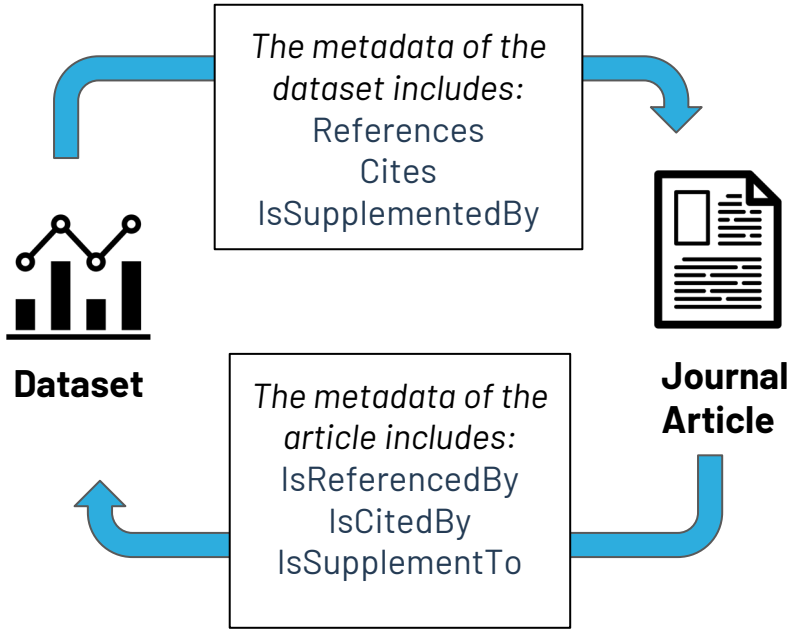


https://support.datacite.org/docs/contributing-citations-and-references

How to create citations



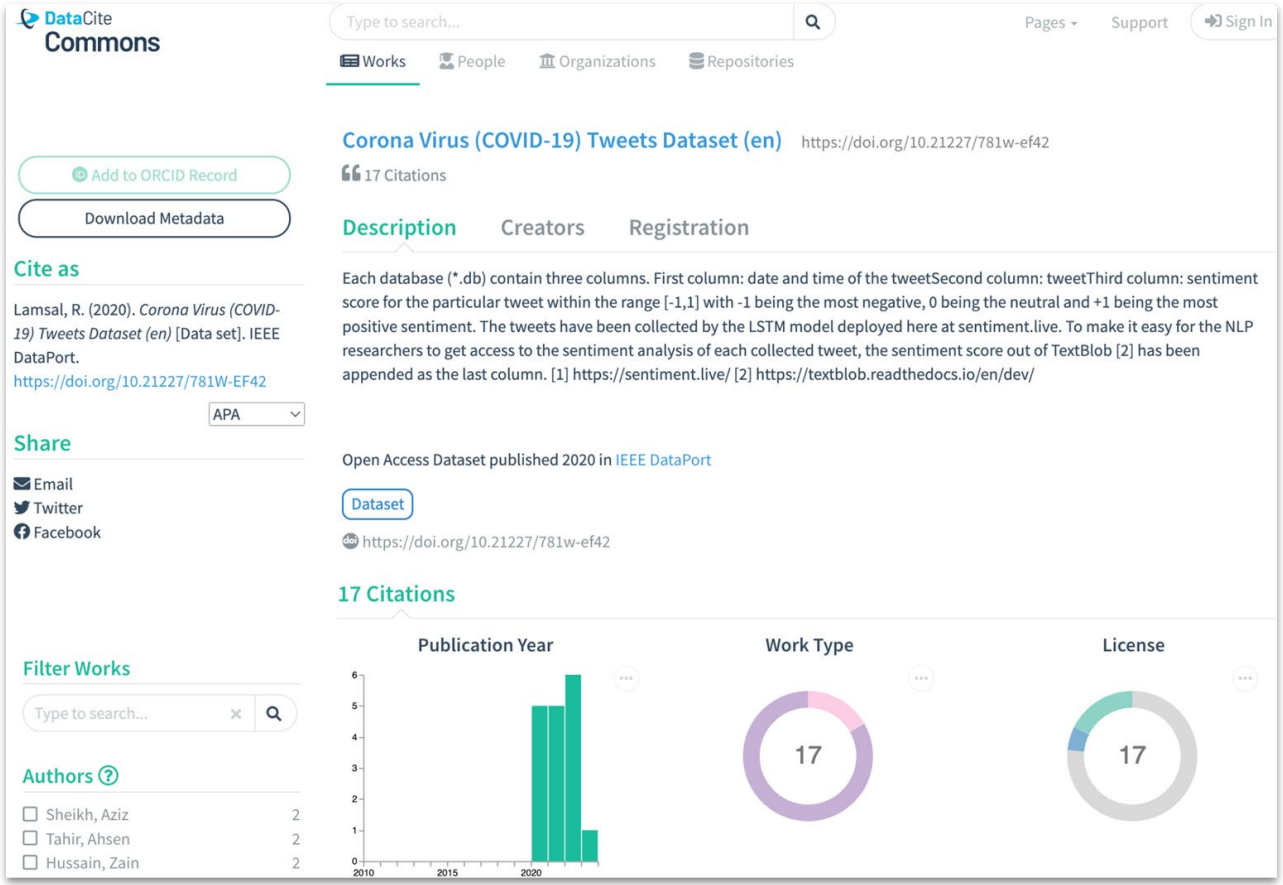
Citation of a dataset



Citation of an article

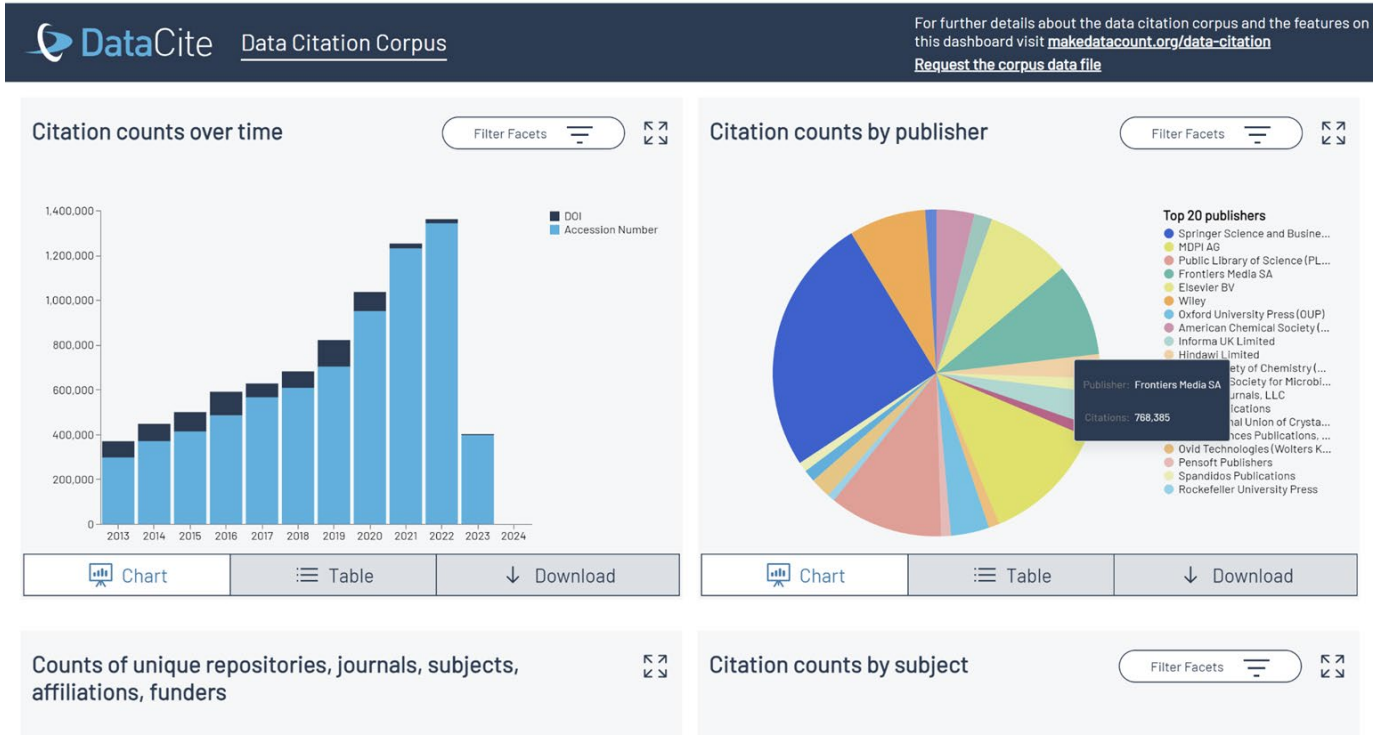
https://support.datacite.org/docs/contributing-citations-and-references

DataCite Commons – surfacing citations



https://commons.datacite.org/doi.org/10.21227/781w-ef42

The Data Citation Corpus



http://corpus.datacite.org/dashboard

World Data System Mission

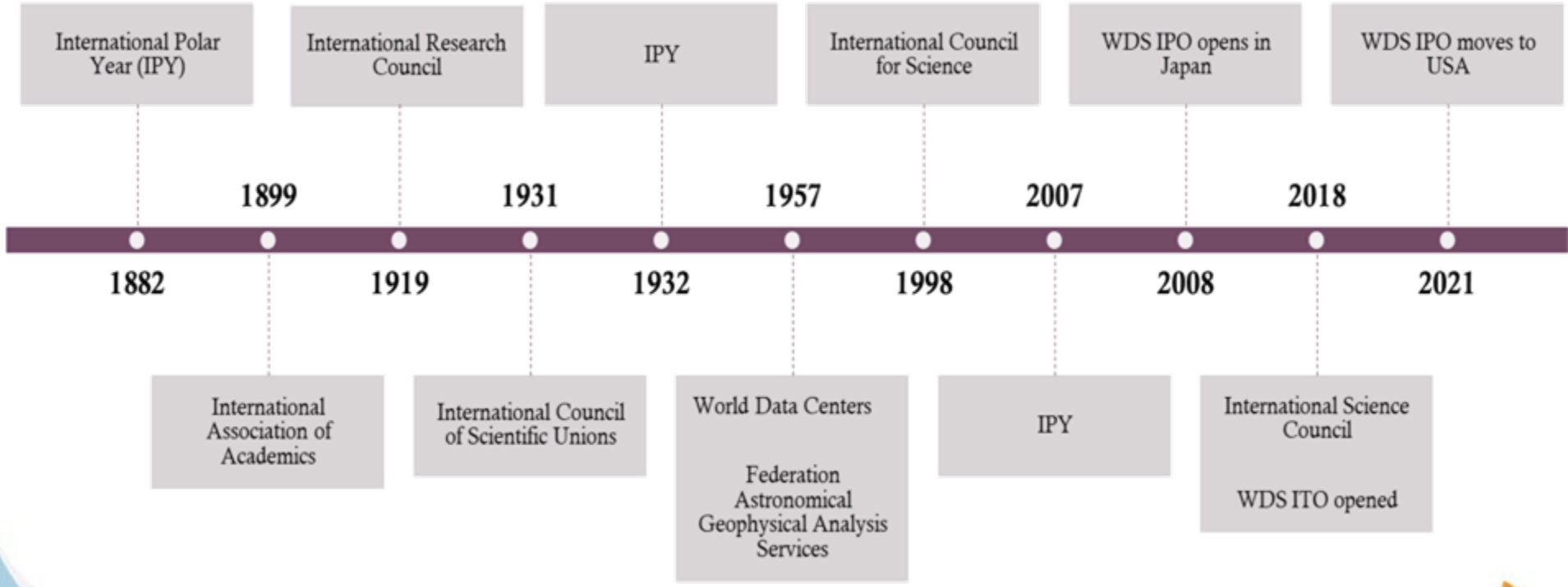
The **mission of the World Data System** is to enhance the capabilities, impact, and sustainability of our member data repositories and data services by:

Creating trusted communities of scientific data repositories

Strengthening the scientific enterprise throughout the entire lifecycle of data and all related components creating first-class data that feeds first-class research output

Advocating for accessible data and transparent and reproducible science

WDS Historical Context



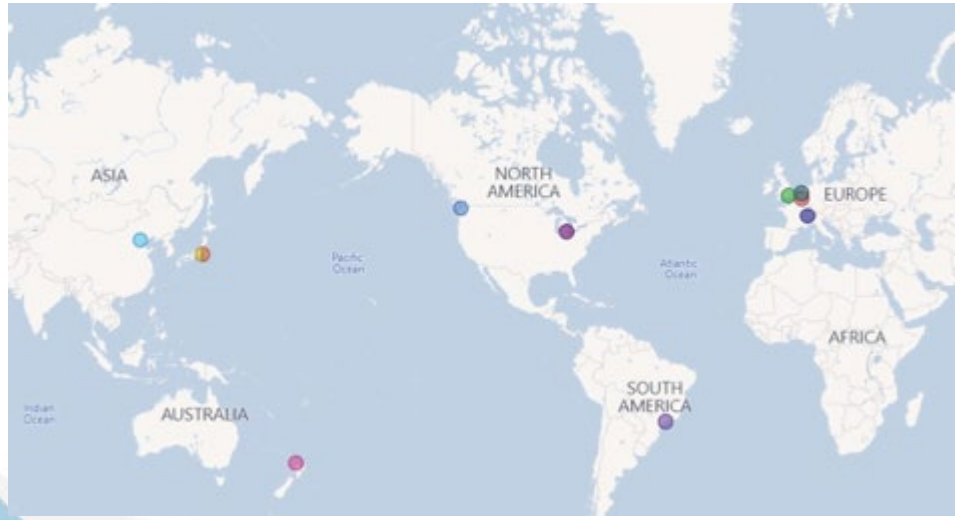
WORLD DATA SYSTEM

WDS Personnel

An affiliate organization of the International Science Council

International Science Council
The global voice for science

The **Scientific Committee** which governs the World Data System is comprised of leading scientists & experts who are actively involved with data; includes directors of WDS member organizations with diverse disciplines & geography.



Scientific Committee
David Castle, Chair

IPO
Meredith Goins
Oak Ridge, Tennessee
USA

ITO
Reyna Jenkyns
Victoria, BC
Canada

WORLD DATA SYSTEM

WDS Membership

91 Regular	Certified Data Repositories
10 Network	Umbrella bodies of data repositories
10 Partner	Contribute support to WDS membership
22 Associate	Interested in the WDS endeavour
Candidate	Preparing or undergoing certification

WORLD DATA SYSTEM

Why WDS endorsed

Aligned with WDS mission, inclusive of many stakeholder types with actionable guidance.

*#1. When publishing their results, researchers deposit related research data and outputs in a **trustworthy data repository** that assigns **persistent identifiers** (DOIs where available). Researchers link to research data using persistent identifiers*

- certification as a basis of WDS regular membership
- CoreTrustSeal Requirements 2023-2025 https://doi.org/10.5281/zenodo.7051012
- R12 - Discovery and Identification: The repository enables users to discover the digital objects and refer to them in a persistent way through proper citation.

Transparency, Responsibility, User focus, Sustainability, Technology

RDA/WDS TRUST Principles Outreach and Adoption Working Group, co-chaired by Meredith Goins

https://www.rd-alliance.org/groups/rdawds-trust-principles-outreach-and-adoption-working-group

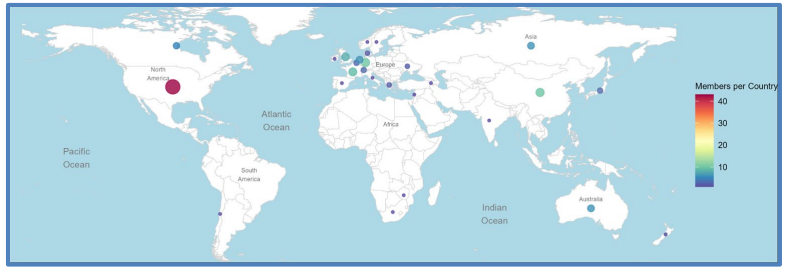
Why WDS endorsed

*#3.Data repositories enable sharing of research outputs in a **FAIR** way, including support for **metadata quality and completeness.***

*#10. Funders, policymaking institutions, publishers and research organizations collaborate towards aligning **FAIR research data policies and guidelines.***

FAIR Principles

- preparing activities for **automated assessments** and FAIRness improvements.
- participant in **FAIR-Impact support action** for "*Recommendations for trustworthy and FAIR-enabling data repositories*"
- PID and licensing support



eosc | FAIR-IMPACT
Expanding FAIR solutions across EOSC

Why WDS endorsed

#4,5,6,7,8,11. Publisher & repository interactions to support data citations, research object relationships and workflows.

Coalition for Publishing Data in the Earth and Space Sciences (COPDESS) is a collaboration among research repositories, scholarly publishers, and other stakeholders focused on leading practices around the preservation and citation of data, software, and physical samples that lead toward credit and reuse.

Newly formed RDA Coordinating Earth, Space, and Environmental Science Data Preservation and Scholarly Publication Processes WG - rationale available for community review until July 12 (login to post comments).

Chairing upcoming conference sessions on PIDs

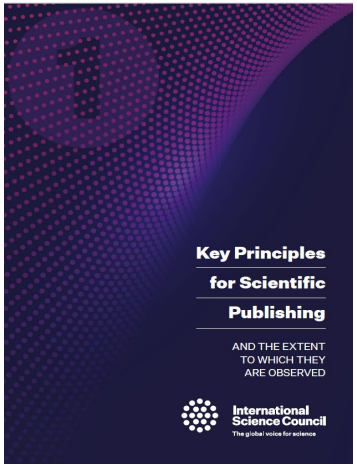
- Earth Science Information Partners Summer Meeting, 2024-07: *Embracing Expanded DataCite Schema Support for IGSNs, Instruments, Publishers, and Funder PIDs*
- American Geophysical Union (AGU), 2024-12: IN019 - *Curating and Interconnecting Persistent Identifiers for Better Research and Funder Mandates* (Abstracts due July 31)

Why WDS endorsed

*#12. Stakeholders responsible for **research assessment** take into account data sharing and data citation in their reward and recognition system structures.*

Research assessment emerged as one of the highest priorities at the 2023 WDS Member Forum.

WDS submitted responses to International Science Council surveys on Research Assessment (2023-12) and Scientific Publishing (2024-02), emphasizing the value of datasets and other research outputs to demonstrate impact.



WORLD DATA SYSTEM

Call to action

Endorse & promote the Joint Statement on Research Data

Share in the chat activities or initiatives that your organization is doing to act upon these recommendations.

Schedule time to take a moment to consider what more you can do.

Discussion

Stay Tuned for Our Riveting Webinar Series!

https://worlddatasystem.org/webinars/

Secure Your Spot in Our Future Webinars

Rewind and Rediscover: Watch Our Past Webinar Sessions

WORLD
DATA SYSTEM

Upcoming Webinar

Navigating the Path to Success in Disaster Data Management

Speakers:

- **Dr. Junshi Xia**, Senior Research Scientist, RIKEN Center for Advanced Intelligence Project, Geoinformatics Team, Japan
- **Dr. Zhang Fen**, Associate Professor, Institute of Geographic Information Science & Provincial Key Lab of GIS at Zhejiang University, P. R. China

Date: 25 June

Time: 01:00 AM UTC

WORLD DATA SYSTEM

For more info:

Email us at wds-ipo@utk.edu

Visit our website at worlddatasystem.org

Instagram: @wds_ipo

Facebook: @WorldDataSystem.International

LinkedIn: @company/world-data-system

X (formerly Twitter): @ISC_WDS

Vimeo: @worlddatasystem

WORLD DATA SYSTEM