Global Summit
Diamond Open Access
BENGALURU 2026

DataCite

CONNECTING RESEARCH,
ADVANCING KNOWLEDGE

From Policy to Infrastructure:
Diamond Open Access Beyond Articles

Mohamad Mostafa
Community Specialist
mohamad.mostafa@datacite.org

3rd Global Summit on Diamond Open Access, Collaboration for Equitable Digital Infrastructures and Knowledge Commons in Scientific Research Systems | 2-6 Feb. 2026 Bengaluru, India

@datacite.org

Bridging Policy → Infrastructure Reality

DataCite

- The journal article is not the research — it's the **final product** of a long research process
- Researchers produce **many outputs and resources**: data, software, models, samples, ETDs, protocols
- These outputs should be **openly shared where possible**, **persistently identified**, and **properly credited** — rather than treated as supplementary or optional
- These outputs are essential for: **transparency • reproducibility • reuse** across the scholarly record
- Yet they **remain largely invisible** in publishing systems & evaluation policies

Making Diamond OA operational for data, code and more

DataCite

Requires a stack of infrastructures, not a single tool:

- Trusted repositories
- Persistent identifiers (PIDs)
- Open metadata standards
- Policies recognising non-article outputs

In 2022, UNESCO published its Recommendation on Open Science, explicitly advocating for increased investment in open research infrastructure and the universal adoption of Persistent Identifiers (PIDs). This framework identifies PIDs, such as DOIs, as essential for the findability of research outputs and resources.

Source: https://doi.org/10.54677/QZPQ1991

unesco

UNESCO OPEN SCIENCE · TOOLKIT GUIDANCE

BOLSTERING OPEN SCIENCE INFRASTRUCTURES FOR ALL

This document is part of the UNESCO Open Science Toolkit, designed to support implementation of the UNESCO Recommendation on Open Science. Building on the provisions of the Recommendation, the guide was developed in consultation with the UNESCO Working Group on Open Science Infrastructures to build a shared understanding and identify steps for strengthening equitable and sustainable open science infrastructures.

What are open science infrastructures?

As defined in the UNESCO Recommendation on Open Science, open science infrastructures refer to shared research infrastructures that are needed to support open science and serve the needs of different communities. These infrastructures may be virtual or physical.

Examples of open science infrastructures include major scientific equipment or sets of instruments, and knowledge-based resources such as collections, journals and open access publication platforms, repositories, archives and scientific data. They also include current research information systems, open bibliometrics and scientometrics systems for assessing and analysing scientific domains, open computational and data manipulation service infrastructures that enable collaborative and multidisciplinary data analysis and digital infrastructures.

The critical components of open science infrastructures allow unambiguous identification of scientific items by unique persistent identifiers[1]. They provide essential open and standardized services to manage and provide access, portability, analysis and federation of data, scientific literature, thematic science priorities or community engagement. These include, *inter alia*, open science platforms and repositories for publications, research data and source codes, software forges and virtual research environments, digital research services and open laboratories.

Different repositories are adapted to the specificity of the items they contain (publications, data or code), to local circumstances, user needs and the requirements of research communities, yet should adopt interoperable standards, diverse practices and best practices to ensure the content in repositories is appropriately vetted, discoverable and reusable by humans and machines.

Additional examples of open science infrastructures that provide common access to physical facilities, capabilities and services include open innovation testbeds, incubators, accessible research facilities, open license stewards, as well as science shops, science museums, science parks and exploratories.

Open science infrastructures are often the result of community-building efforts, which are crucial for their long-term sustainability and therefore should be not-for-profit and guarantee permanent and unrestricted access to the public to the greatest extent possible.

Infrastructures are key to the sustainability of open science

In the Recommendation, open science infrastructures are defined as a pillar of open science, alongside open access to scientific knowledge, open engagement of societal actors and open dialogue with other knowledge systems.

For maximum efficiency and impact, open science should build on long-term practices, services, infrastructures and funding models that ensure the equal participation of scientific producers from less privileged institutions and countries.

Open science infrastructures should be organized and financed on a primarily not-for-profit and long-term vision, which enhance open science practices and guarantee permanent and unrestricted access to all, to the largest extent possible.

Examples of unique persistent identifiers (PIDs) include ORCID IDs for people, digital object identifiers (DOIs) for publications and grants, DataCite DOIs for datasets, ROR IDs for organizations, research activity identifiers (RAID) for research projects

Global Reality

DataCite

In well-resourced regions:

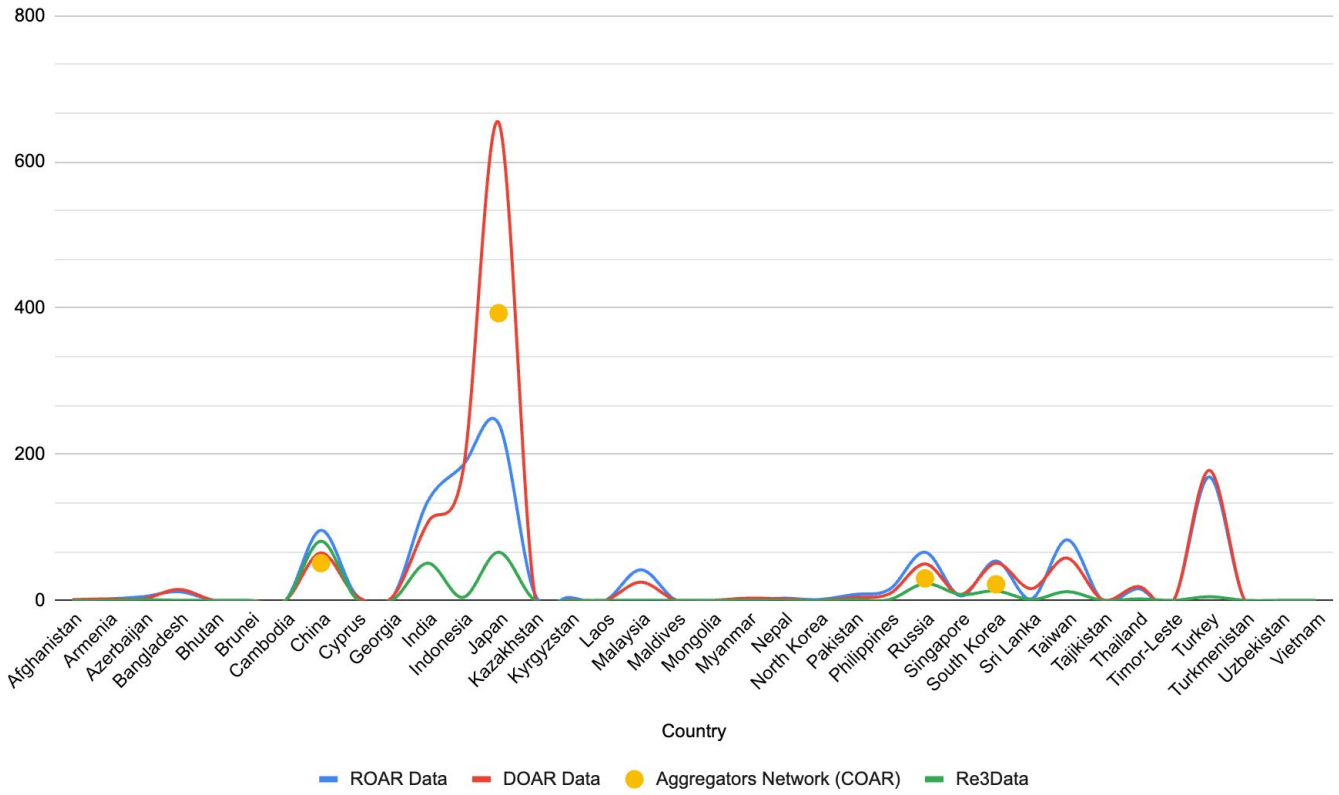
- PIDs are widely adopted.
- Metadata flows across systems
- Outputs linked to funders & institutions

Emerging regions face:

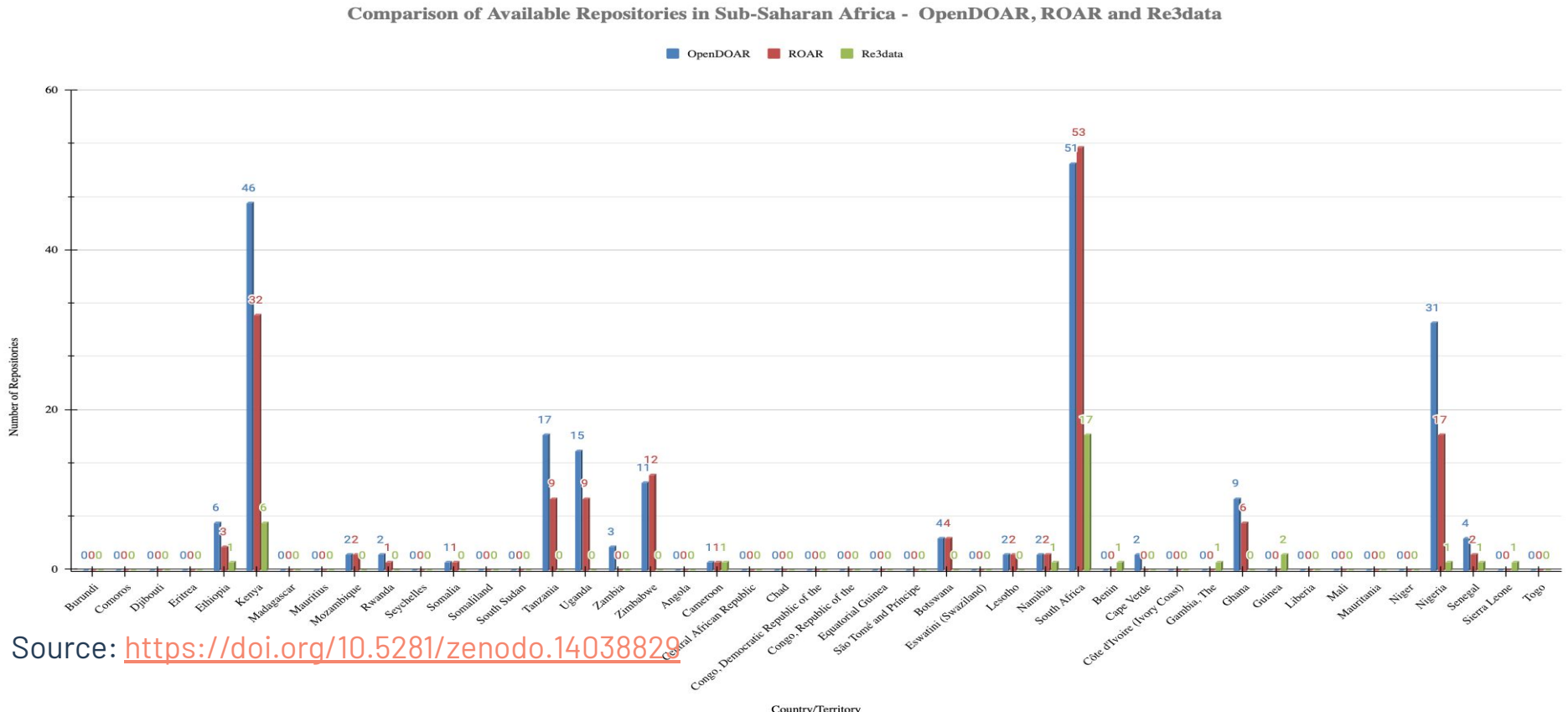
- Lack of underlying infrastructure (data repositories)
- Limited metadata capacity
- Weak or absent national Open Science policies
- No sustained funding for infrastructure

OA Repositories infrastructure in Asia

Source: https://doi.org/10.5281/zenodo.12566244



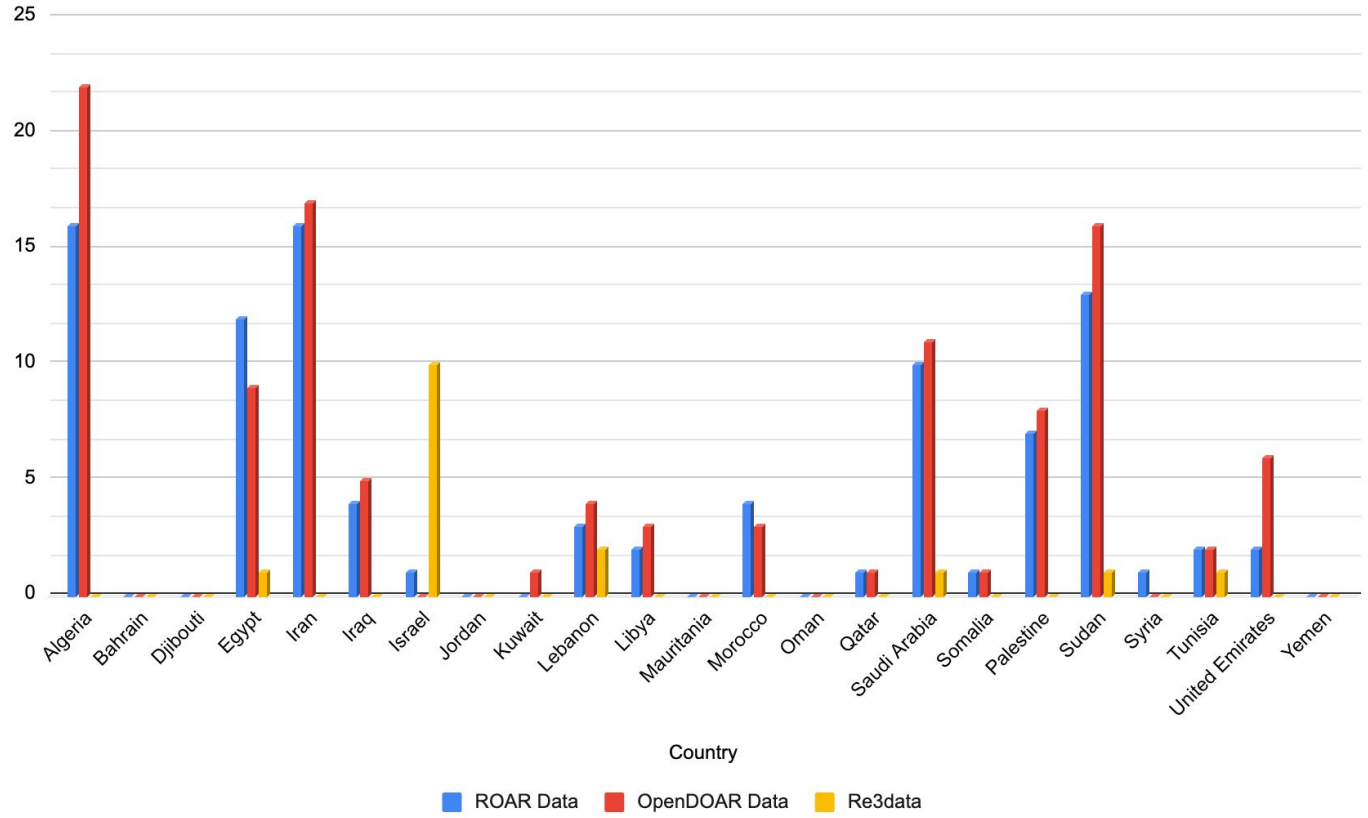
OA Repositories infrastructure in Sub-Saharan Africa



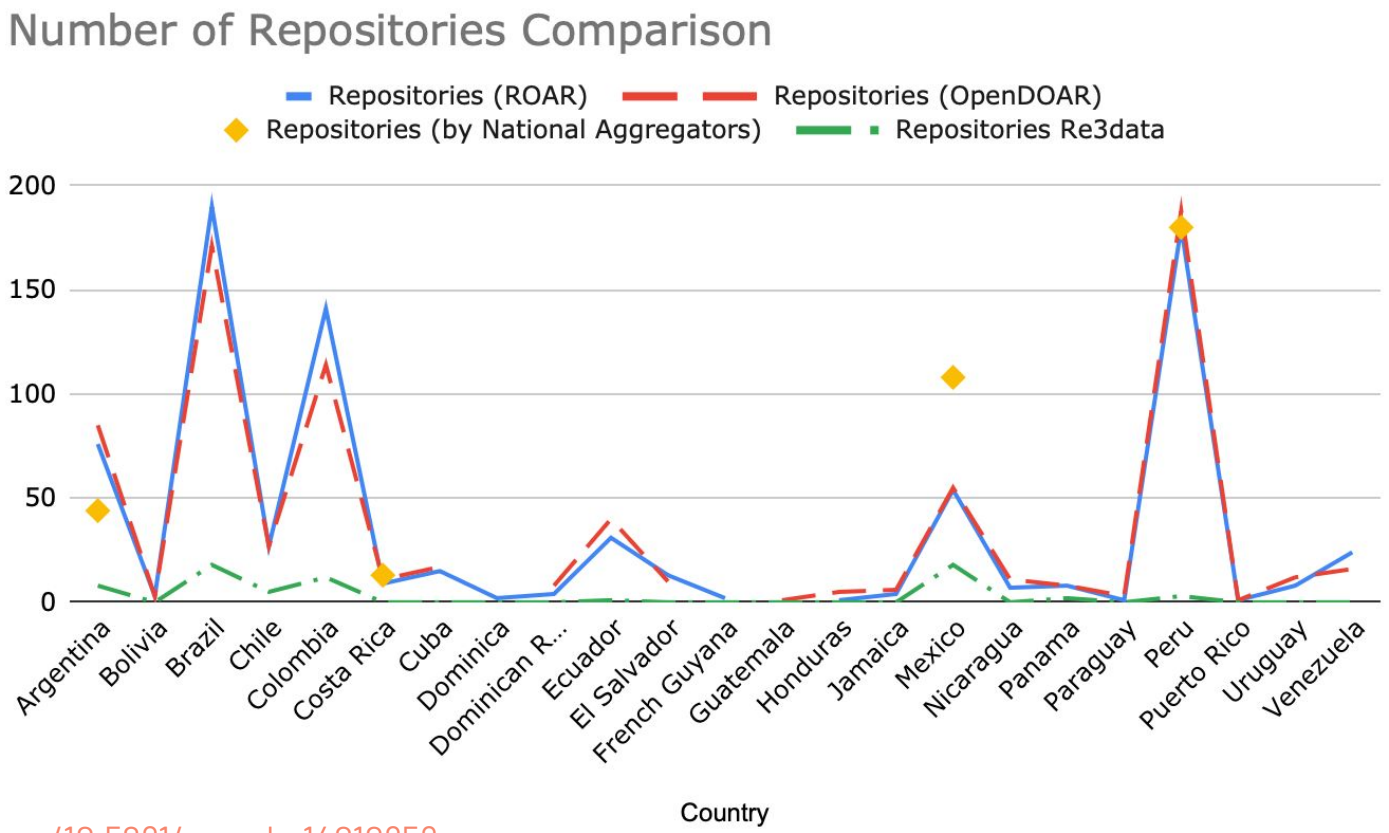
Source: https://doi.org/10.5281/zenodo.14038829

OA Repositories infrastructure in MENA

Source: https://doi.org/10.5281/zenodo.11370031



OA Repositories infrastructure in Latin America



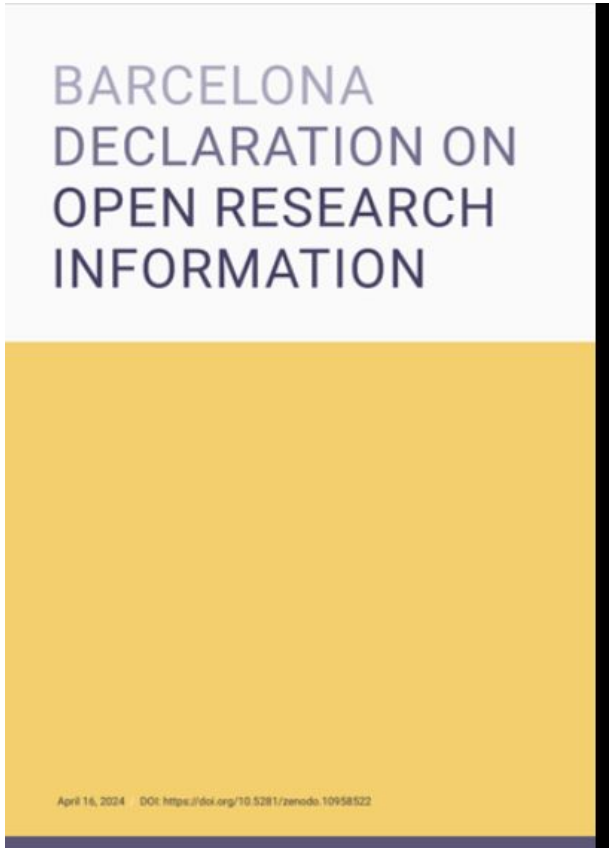
Source: https://doi.org/10.5281/zenodo.14010858

Why Policy Matters

Why Global Policy Matters

- UNESCO Open Science Recommendation & the Barcelona Declaration:
  - Identifies open research infrastructure as a public good

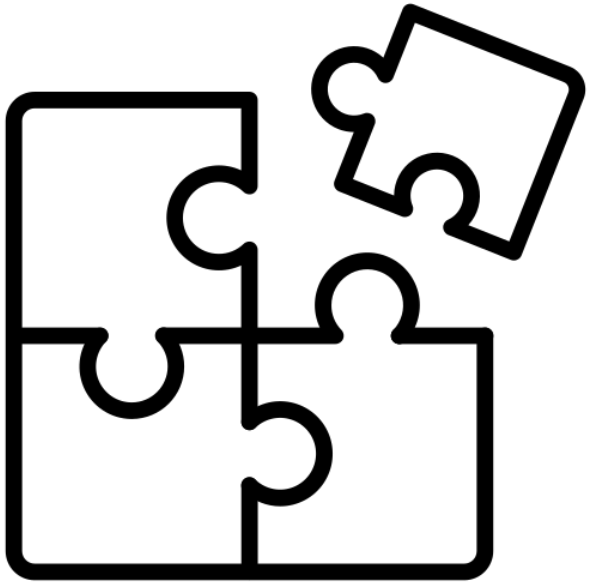
- Without sustained investment:
  - Diamond OA becomes structurally exclusive
  - Visibility and credit concentrate in a few regions



Role of PIDs & Metadata

The invisible infrastructure that makes OA work

- PIDs (e.g. DOIs) enable:
  - Findability
  - Citation & credit
  - Linking across research workflows
- Metadata provides:
  - Context & relationships
  - Connections to people, funders, institutions
  - Reuse, replication, validation



What Breaks Without Them

Data exists but isn't discoverable; Reuse happens without attribution; Outputs fall outside evaluation systems; Repositories/research system become non-interoperable

Risk of:

Fragmentation; Two-tier global research visibility; Dependence on closed systems

Data Exists, Recognition Doesn't

The evaluation gap

- Technically:
  - DataCite provides infrastructure to assign DOIs to more than 30 output types
  - Cite data and software
  - Link outputs via metadata

- But institutions still: Prioritise articles; Ignore data & software citations

CoARA is working to reform research assessment by moving beyond journal-based metrics and explicitly recognising diverse research outputs — including data, software, and other non-article contributions.

About DataCite

DataCite

We are a global community of research organisations sharing 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.

A **non profit organization** registered in Hannover, Germany since 2009.

Our Community

DataCite

1737

Organizations

68

Countries

113m+

DOIs

1B+

Resolutions in 2025

(Feb. 2026)

Our services

DataCite

Create DOIs

Enable the creation and management of persistent identifiers

Connect research

With rich metadata for diverse outputs – from samples and images to data and preprints

Integrate services

to improve and enhance research workflows

Facilitate discovery and reuse

of research outputs and resources with flexible state-of-the-art tools and technology

Resource Types in DataCite Fabrica

DataCite

Dataset	61,259,169
Text	17,339,953
Physical Object	16,960,491
Image	10,557,607
Other	3,559,281
Journal Article	2,930,174
Preprint	1,804,588
Collection	1,203,690
Software	813,041
Audiovisual	661,711
Dissertation	525,079
Conference Paper	346,317
Report	318,645
Study Registration	216,164
Book	201,160
Interactive Resource	178,627
Project	142,852
Book Chapter	132,849
Event	129,505
Journal	88,984
Sound	80,434
Model	45,833
Conference Proceeding	42,620
Peer Review	32,611
Data Paper	18,643
Standard	11,879
Computational Notebook	11,727
Workflow	10,789
Output Management Plan	8,793
Award	8,115
Instrument	1,227
Service	570

(Feb. 2026)

Key messages: DataCite & Diamond OA

- DataCite's mission is to **connect research and advance knowledge** by providing **persistent identifiers (DOIs) and open metadata** that support a fair, sustainable, and community-driven scholarly ecosystem—fully aligned with the principles of Diamond Open Access.
- Beyond articles, **DataCite DOIs can be assigned to more than 30 types of research outputs, resources, activities**, including datasets, theses, dissertations, reports, software, awards, and more—supporting the full diversity of scholarly contributions central to Diamond OA.
- Through **open, rich metadata and a robust schema**, DataCite enables **meaningful relationships between research outputs** (linking articles to data, software, funders, institutions, and infrastructures), strengthening interoperability and reuse across the research landscape.
- DataCite is a **global, community-led infrastructure**, bringing together universities, libraries, funders, and research organisations—many of whom are directly engaged in Diamond Open Access initiatives and non-commercial publishing models.
- More broadly, **DataCite acts as an Open Science enabler**, supporting transparency, openness, and collaboration across research systems—making it an essential infrastructure partner for institutions and communities advancing Diamond Open Access worldwide.

Diamond OA beyond articles

DataCite

Diamond Open Access is **not just about removing paywalls from articles**. It's about **building the infrastructure that allows all forms of knowledge** — data, software, methods, and more — to be **visible, reusable, and valued**.

Infrastructure → lived ECR reality

All of this infrastructure discussion **ultimately lands on people** — particularly early-career researchers — because they are the ones producing data, software, and metadata, **often without clear recognition or reward.**

DataCite

CONNECTING RESEARCH,
ADVANCING KNOWLEDGE


info@datacite.org


pidforum.org


datacite.org
datacite.org/blog


support.datacite.org
support@datacite.org


@datacite


@datacite


@datacite


DataCite