Working with DataCite

Identify, cite, and connect research with confidence.

About DataCite

WHO WE ARE

DataCite is a leading global non-profit organization that provides persistent identifiers (specifically DOIs) for research data and other research outputs. Organizations within the research community join DataCite as members to assign DOIs to all their research outputs. This allows their outputs to become discoverable, citable, and connected. Associated metadata is made available to the community and DataCite develops additional services to connect and share research outputs with the broader research ecosystem.

OUR VISION AND MISSION

Connecting research, Identifying knowledge.

DataCite's mission is to be the world's leading provider of persistent identifiers for research. Through our portfolio of services, we provide the means to create, find, cite, connect, and use research. We seek to create value and develop community-driven, innovative, open, integrated, usable, and sustainable services for research.

WHAT WE DO

DataCite works primarily with repositories and the organizations that manage them. DataCite DOIs provide a solution that enables easy identification of the research deposited in repositories. When a DOI is assigned to any research output, it not only becomes identifiable, but through the associated metadata and central search index, it can be discovered and connected to other research outputs.

HOW WE WORK

We are community owned and driven and our structure allows us to be proactive, agile, and responsive to the needs of the research community. Our guiding values are trust, transparency, and openness. All DataCite's metadata is open, and all software is open source. We advocate for the role of all research content in the research landscape – we engage in outreach that reflects the interests of our diverse community.

WHY WE DO IT

Research outputs need to be shared as openly as possible, but they also need to be made available in a way that makes sense to both humans and machines. The FAIR data principles lay out concrete steps to make data Findable, Accessable, Interoperable, and Reusable and consolidate the role of Persistent Identifiers (PIDs) in achieving this. By providing reliable community owned infrastructure to register DOIs and metadata, DataCite plays a key role in making data FAIR.

Getting started

1

Decide if DOIs are right for your organization.

There are certain responsibilities associated with registering DOIs you should consider (see below).

2

Complete the member form to express your interest.

Express your interest by completing the form on our website and we'll work with you to understand your needs and share information about our services.

3

Submit your membership application.

Your application will be evaluated and approved by the DataCite members.

4

Schedule an onboarding session.

Join an onboarding session to learn everything you need to know about DataCite.

5

You're officially part of DataCite.

You're officially part of DataCite! Now you can register DOIs and share your research in a FAIR compliant way.

DataCite membership

DataCite is open to all organizations that share our commitment to sharing research. You can join DataCite if you want to register DOIs and use our services, but also to support our mission and be part of our governance. Among our members you'll find academic organizations, research institutions, publishers, government agencies, funding agencies, societies, and infrastructure providers.

If you want to join DataCite for DOI registration purposes, it is important to consider the following questions:

1. Do you have the authority to assign DOIs to research outputs?

2. Will you be able to guarantee persistence?

3. Will you be able to provide and maintain metadata for each resource with a DOI?

4. Will you be making metadata openly available without restriction?

5. Are you planning to make the data accessible to external users where possible?

6. Are the research outputs ready to be cited by others?

7. Will you maintain publicly accessible landing pages for each item with a DOI?

If the answer to all of these questions is yes, you can start registering DOIs via DataCite.

There are different membership options if you want to join DataCite, dependent on your needs:

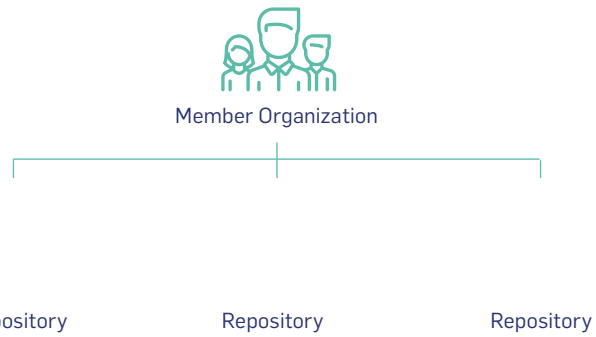
MEMBER-ONLY

This type of member supports DataCite's mission and wants to collaborate with DataCite and/or be part of DataCite's governance. These members do not register DOIs.

Member Organization

DIRECT MEMBER

This type of member is a single organization that joins DataCite as a member in its own right and consumes DOI services. This single organization may have one or more repositories under their umbrella, but those repositories are under the same administrative structure as the organization. A direct member also takes on membership responsibilities such as participating in DataCite's governance.



CONSOR

ations that have come together to collectively participate
ties and use DataCite's DOI services. A consortium is
ons that are under different administrative structures.
try or subject based. Organizations within a consortium can work with one or more repositories. One organization takes on the role of Consortium Lead.

Organizations within a DataCite consortium can take advantage of a consortium arrangement for a variety of reasons:

1. Organizations may not have the resources or capability to join individually.
This is particularly true for small organizations.

2. Organizations that are already collaborating, are working within the same discipline,
or within the same (language) area may benefit when developing a shared PID strategy.

3. Consortia can leverage relevant skills, know-how, and expertise of each organization within the consortium. This can increase the uptake of DOI services without overstretching the resources of any single organization.

4. A consortium is able to speak with one voice.

Repositories

DataCite's core activity is providing tools for DOI and metadata registration for repositories to make data FAIR. Therefore, repositories play a key role in the DataCite member model. We define a repository as a service operated by research organizations, where research materials are stored, managed and made accessible. A repository is a single unit and DataCite links the repository to information in re3data, where additional repository metadata are available.

However, we realize that not all content is hosted in repositories and therefore DataCite also includes periodicals as a designation. The term periodicals includes journals, proceedings, books, blogs, and working paper series.

The repository information provided will be used and displayed throughout DataCite services including search, data usage statistics, member DOI statistics, and third-party indexing services.

Member roles and responsibilities

There are different roles and responsibilities associated with each member type.

	Member responsibilities	Service responsibilities	Coordination responsibilities
Member only	✓		
Direct member	✓	✓	
Consortium lead	✓	✓	✓
Consortium organization		✓	

DataCite

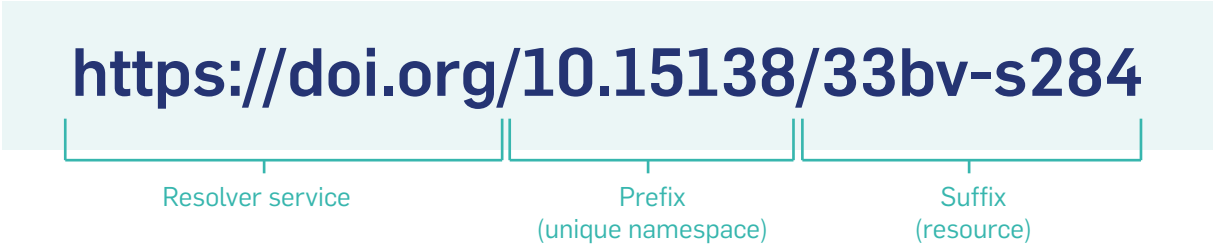
What is a DOI?

DOI is an acronym for “digital object identifier”. DOIs are a type of Persistent Identifier (PID) that uniquely identify digital research content. They are intended to be a permanent way of identifying and accessing a particular resource. DOIs form a persistent link that points to the repository or other digital location by including the URL in the metadata. This provides a system for persistent and actionable identification and interoperable exchange. DOIs remain fixed, but the location and other metadata may change. DataCite DOIs come with a metadata schema that includes a controlled vocabulary of 15 different resource types to describe the content being shared. DataCite members are responsible for updating and managing their DOIs and metadata. DataCite provides a number of services around DOI registration to maximize the benefits of DOI use.

DOIs ARE

1. **Well-established as identifiers for research outputs**
2. **Recognized as an ISO standard**
3. **Centrally managed and governed by the International DOI Foundation (IDF)**
4. **A system for persistent and actionable identification.**
5. **Widely used by the research community**

EXAMPLE OF A DOI



A DOI name is made up of a prefix and a suffix separated by a forward slash. The prefix is a "namespace" used to ensure that DOIs are globally unique. A DOI prefix always starts with '10.' and continues with a number (e.g. '10.1234' or '10.20865'). Each DataCite member organization is assigned its own unique prefix(es). While members can decide their own suffix format, DataCite recommends random opaque strings.

A DOI is always associated with a metadata description of the object as well as to a digital location, such as a URL directing to a landing page, where all the details about the object are accessible.

DATACITE DOI s ARE SUITABLE FOR A WIDE RANGE OF RESEARCH OUTPUTS

1. Research datasets and collections, associated workflows, software, images, and models

2. Grey literature such as theses, dissertations, reports, unpublished conference papers, newsletters, preprint journal articles, technical standards, and specifications for which the institutional repository is the primary publication point.

Registering DOIs

DataCite offers its members two methods to register DOIs: the Application Programming Interface (API) and DOI Fabrica, our web interface. Once you become a member you will be assigned an account and be able to access both the API and DOI Fabrica.

(API)
APPLICATION PROGRAMMING INTERFACE

The DataCite REST API allows DataCite members to register DOIs and manage their repositories and prefixes. Organizations generally integrate with the API to provide their own in-house DOI registration services for their repository users. This option is preferable for registering large numbers of DOIs but requires technical know-how. The DataCite REST and OAI-PMH APIs allow any user to retrieve, query and browse DataCite DOI metadata records.

WEB INTERFACE

DOI Fabrica is a one-stop service to create and manage DOIs for your organization. It is also where you can manage your DataCite account contact information and associated repositories and prefixes. The user-friendly web interface allows you to enter the metadata in a web form or upload them via a file in one of the supported metadata formats. Members who register smaller numbers of DOIs like to use the web interface.

Metadata

The DataCite Metadata Schema is a list of core elements which were chosen to support the citation and description of research outputs. As a DataCite member, you can use the DataCite Metadata schema to describe the objects that you are sharing. The key to making data citable, searchable and accessible is enriching datasets with metadata descriptions about the data. The DataCite Metadata Working Group develops the DataCite Metadata Schema and its documentation, in consultation with DataCite members and staff.

SOME KEY BENEFITS
OF THE DATACITE METADATA SCHEMA ARE:

It forms the basis for a **recommended standard citation format** for research.

It is **interoperable with other data management schemas**, such as Dublin Core.

It can accommodate **numerous types of research outputs.**

Its development is overseen by a Working Group composed of DataCite members, allowing the schema to **evolve in response to feedback** from DataCite members.

The Metadata Schema consists of six mandatory properties;
Creator, Title, Publisher, Publication Year, Resource Type, and Identifier
These six elements support data citation. There is also a range of recommended and optional properties that may be used to describe and identify the objects.

Connecting data

Using persistent identifiers becomes truly powerful when these identifiers are connected to each other. DataCite helps you connect your research data to other research outputs in several ways:

1.ORCID AUTO-UPDATE

As an ORCID member and trusted partner, DataCite can ask researchers for permission to automatically update their ORCID record. Once this is enabled, whenever a DataCite member sends us metadata that includes an ORCID iD, then we will automatically add it to their ORCID record.

2. DATA CITATION & EVENT DATA

Data citation establishes links between articles and datasets. This increases transparency and allows researchers to get credit for their work. By including information about related identifiers in their metadata, DataCite member repositories can indicate how their datasets have been cited. This information then feeds into Event Data, a joint service by DataCite and Crossref.

3. PID GRAPH

DataCite is working with its partners in the FREYA project to develop the PID Graph, a network of interconnected PID systems, as a basis for a wide range of services. The PID graph can link PIDs together via relations in their metadata to enable the discovery of connections two or more steps away. This allows us to e.g. aggregate the citations for all datasets hosted in a particular repository, funded by a particular funder, or created by a particular researcher. Joining DataCite means that your datasets automatically become part of the PID Graph and are connected to other research outputs, researchers, institutions, and funders.

Tracking data use

DataCite DOIs make your research data and other outputs discoverable and citable. This means it becomes easier for the data to be reused. DataCite wants to help you track reuse of your data. We, therefore, make data citations obtained via our collaborator, Crossref, available through our API and in our search index.

DataCite was also one of the collaborators involved in writing the COUNTER code of practice for research data and works with members to help them count views and downloads of their datasets. When these counts are submitted to DataCite, DataCite makes these available through APIs and in our search index, making it easy for you to show the world how your data are being reused.

Downstream impact

If you are putting resources into sharing your research, you want it to have impact. It's therefore important that DataCite metadata is available to everyone.

To showcase and expose metadata, DataCite provides an integrated search interface, DataCite Search, where you can search, filter, and extract all the details from a collection of millions of records. This interface is complemented by our REST and OAI-PMH APIs, providing a machine actionable end-point to harvest DataCite's metadata collection.

DataCite Search is used by Google Dataset Search to index all research data with a DataCite DOI. That means that if you are registering DataCite DOIs, your research data will automatically appear in Google Dataset Search, thereby reaching a wide audience.

DataCite metadata are also harvested by organizations such as OpenAIRE and Clarivate's Data Citation Index

DataCite community

SUPPORT

DataCite provides its members with tailored support via our support desk support@datacite.org. No problem is too big or small for our support team. We like to provide the best user experience possible for our members by responding in a timely way, listening to their feedback, and maintaining our support site with guides and documentation to help make navigating the world of DOI registration easy.

PIDFORUM

The PIDforum.org is a discussion forum for all things related to PIDs. DataCite members have access to a private DataCite Community group, which includes a Chat Room where you can discuss your user stories, talk to other members and hear the latest news and views from the DataCite community.

COMMUNICATION CHANNELS

DataCite believes that communicating and engaging with our community is a key part of our mission. Next to our website, the regularly updated DataCite blog provides readers with news of all the latest product releases, conferences and other interesting updates. To help keep our community informed we also send out a monthly newsletter and regularly update Twitter and Youtube.

THE ROADMAP

The DataCite roadmap is informed by its members. The development of new services and features is carefully mapped based on feedback. We make a public roadmap available on our website which allows users to upvote, so it's even easier for you to add your ideas and suggestions to help guide what we build and where we innovate.

OPEN SCIENCE COMMUNITY

DataCite is active in the Open Science community, attending numerous conferences and participating in working groups every year to make sure we keep up to date with the latest in innovation, technology, and best practice. These include: RDA, Open Repositories and Force11 among others. DataCite is one of the organizers of PIDapalooza, a festival for persistent identifiers. DataCite is also a trusted partner in numerous publicly funded projects and plays an active role in the development of the European Open Science Cloud (EOSC).

Stay in touch!

Email us:
info@datacite.org

Follow us:
@datacite

Talk to us:
pidforum.org

Website:
datacite.org

Working with DataCite

Identify, cite, and connect research with confidence.

www.datacite.org

DataCite
FIND, ACCESS, AND REUSE DATA