DataCite

CONNECTING RESEARCH,
ADVANCING KNOWLEDGE

DataCite – Contributing to PIDs infrastructure in Africa

Bosun Obileye
Regional Engagement Specialist, Africa

30th of November 2023
AfricArxiv Webinar Series

@datacite

@datacite@openbiblio.social
@paul4kant@scicomm.xyz

CC BY

DataCite

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 community, we make research more effective with metadata that connects research outputs and resources–from samples and images to data and preprints. We enable the creation and management of persistent identifiers (PIDs), integrate services to improve research workflows, and facilitate the discovery and reuse of research outputs and resources.

Our community

DataCite

3000+

Repositories

680+

Members

51

Countries

56m+

DOIs

1400+

Organizations

(July 2023)

PIDs?

Menti.com 6736 2074

DataCite

or
Scan the QR code

Or
https://www.menti.com/alvi4fc5p9fp

What are PIDs?

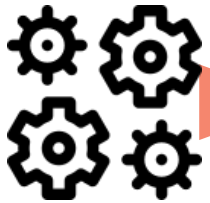
DataCite

- PIDs implies Persistent Identifiers
- PIDs are labels that locate, identify and share information about digital objects.
- They are used for ***identifying resources***. For instance, a journal article, preprints, conference proceedings, books, book chapter, datasets, samples, archaeological outputs, institutions, individuals, and many more.
- PIDs must ***be persistent***. A PID, once created, should always exist. It should never be deleted. It is possible to update what a PID points to or to update the metadata but the identifier must remain constant.
-
- A PID may be connected to a set of metadata describing an object rather than to the object itself.

PIDs

What is a persistent identifier (PID)?

https://doi.org/10.34848/GJO6SY



https://research-data.urosario.edu.co/dataset.xhtml?persistentId=doi:10.34848/GJO6SY

Unique alphanumeric string referring to a digital resource.

Always points to the same resource (a metadata representation)

DOIs for research outputs and resources
https://doi.org/10.5281/zenodo.3630248

ORCID iDs for researchers
https://orcid.org/0000-0001-6622-4910

ROR IDs for research organizations
https://ror.org/01y2jtd41

DataCite

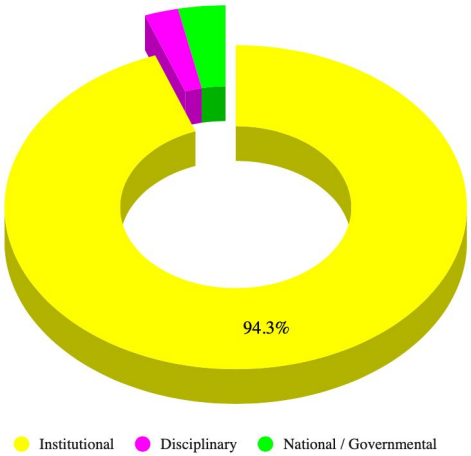
PIDs infrastructure in Africa

PIDs landscape
Do you know that?

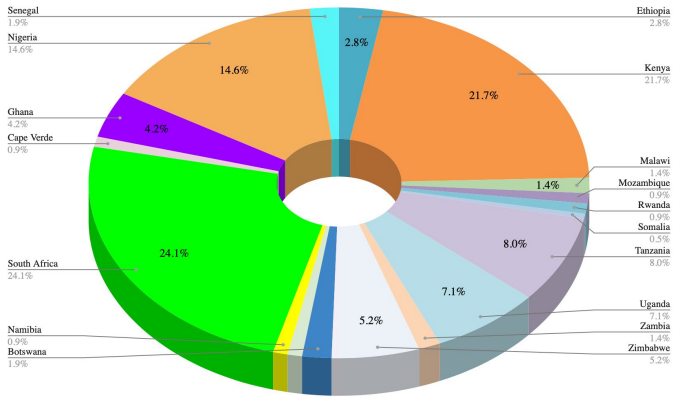
According to OpenDOAR, about 94% of repositories in Africa are Institutional

About a quarter of repositories in Africa are located in South Africa while Kenya holds about a fifth of the total

Type of Repositories in Sub-Saharan Africa.

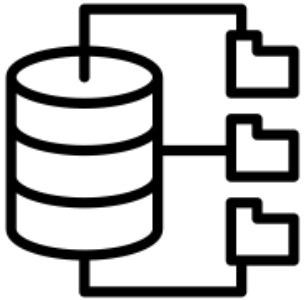


Number of Repositories Per Country. source - OpenDOAR

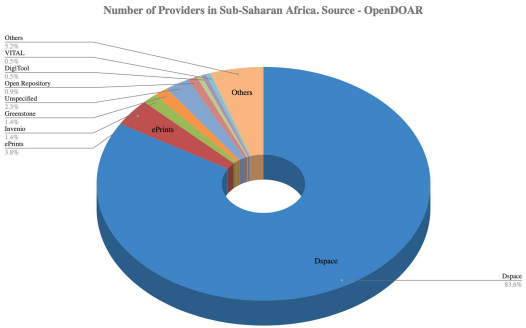


Do you know that...

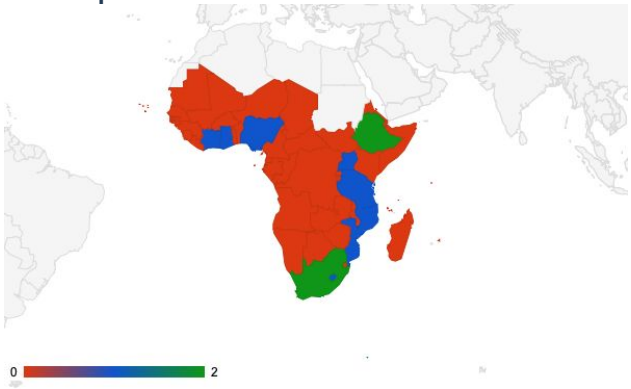
Africa has about 100,000 findable DOIs in DataCite's Commons



Dspace accounts for more than 80% of repositories software in Africa



According to UNESCO only 2 countries has implemented open science in Africa



What could be the challenge(s)?

Let’s go to menti.com

Menti.com 6736 2074

DataCite

or
Scan the QR code

Or
https://www.menti.com/alvi4fc5p9fp

Challenges

1. Awareness - Low awareness about the need for PIDs
2. Institutional support - Poor National and institutional policy on open science
3. Funding - Poor funding or lack of funding
4. Knowledge gap - technical knowledge and skills
5. Mindset on digitization and open access
6. Ethical fears on reuse, license, Intellectual Property Rights
7. Human resources - getting the right persons for the job

Opportunity for growth

1.Awareness about PIDs is gaining momentum in the region

2. New entrants are exploring other repository software

3. Policy makers are warming towards open science

4. Funders are pushing for use PIDs in open science/open access

5. Digital publications is gradually replacing paper based publications.

6. Knowledge preservation is stimulated with digitization using PIDs

DataCite

Strategic Initiative

Strategic Initiatives
It Takes a Village

In line with our mission and vision, DataCite also actively participates and leads various initiatives through collaboration with stakeholders in the community to make open science a reality.

- **Data metrics** – We help further the adoption and implementation of responsible data metrics with, for example, the Make Data Count initiative.
- **Identifier registries** – We support community-led registries of identifiers such as the Research Organization Registry (ROR).
- **Repository discovery** – We contribute to the development of repository discovery initiatives such as re3data with collaboration and financial support.

https://makedatacount.org/

https://ror.org/

https://www.re3data.org/

The Global Access Program (GAP)

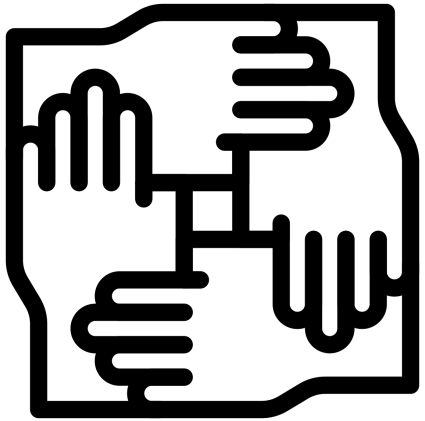
Global Access Program

DataCite

- Launched in January 2023
- Funding from the Chan Zuckerberg Initiative
- Improve equity and access to our PID infrastructure in underrepresented regions through a **comprehensive** approach

Towards an Equitable and Inclusive Research Ecosystem

The goal of the DataCite Global Access Program is to help to **create a more equitable and inclusive research ecosystem**, where all researchers and communities have the tools and resources they need to conduct and share their work.



The Team

Gabriela Mejias

DATACITE'S COMMUNITY & PROGRAM MANAGER

In her role, she leads the Global Access Program, contributes to DataCite's outreach efforts and seeks collaboration with the research community. She also leads DataCite participation in the FAIR-IMPACT project. Previously, she worked at ORCID focusing on community engagement, to increase adoption and membership. Gabriela volunteers at the EOSC PID Policy & Implementation Task Force, the Board of Directors of the Networked Digital Library of Theses and Dissertations and the NISO Diversity Equity Inclusion and Accessibility committee. She's interested in shaping a more open research infrastructure.

Mohamad Mostafa

REGIONAL ENGAGEMENT SPECIALIST, MIDDLE EAST AND ASIA

Mohamad joined DataCite in June 2023 as a Regional Engagement Specialist for the Middle East and Asia. He works with the community to build more openness and trust in scholarly infrastructure and supports emerging communities transition towards Open Research and implementing its principles. Mohamad has participated in launching the ORCID Arabic interface to the Arab world and served as a volunteer Crossref Ambassador for the MENA region. Mohamad is based in Dubai (UAE) and he is passionate about Open Science and has been raising awareness among the research community.

Bosun Obileye

REGIONAL ENGAGEMENT SPECIALIST, AFRICA

Bosun spent the past 6 years before joining DataCite serving as the Institutional Data Manager for International Institute of Tropical Agriculture, IITA, having presence in more than 35 African countries. He was instrumental to the institutionalization of Open Data in the organization. Within CGIAR, he was active in various Community of Practices (CoP) and working groups like Repository Working Group, Metadata Working Group,among others. He had earlier spent his immediate years as a leader in the Information Technology industry. Bosun spends his leisure time with his family, playing the piano, reading, editing videos and walking.

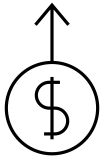
Arturo Garduño-Magaña

REGIONAL ENGAGEMENT SPECIALIST, LATIN AMERICA

Arturo is a passionate advocate for open access and equitable research practices who joined the DataCite team as the Regional Engagement Specialist for Latin America. He is responsible for promoting and implementing the program in the region, with the aim of enhancing equitable access to infrastructure and fostering the adoption of persistent identifiers.

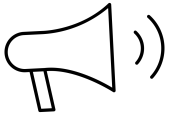
Global Access Program
Three Components

DataCite



Funding. The program will provide the opportunity for communities to seek funding to support activities related to the program.

Technical infrastructure. Through partnerships with local and international stakeholders, the program will seek to support communities in building technical infrastructure per their needs (and design).

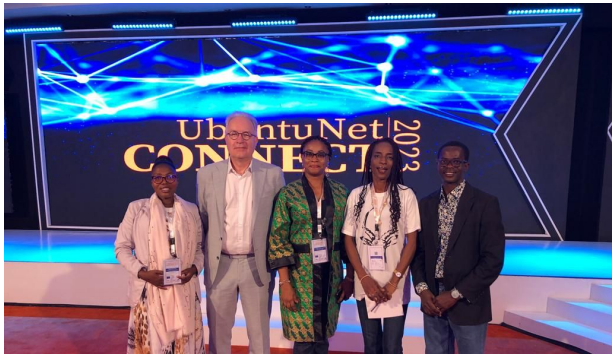


Outreach. In learning from local community stakeholders, we seek to increase awareness of PID infrastructure.

Outreaches

DataCite

IASSIST Workshop



GAP webinar

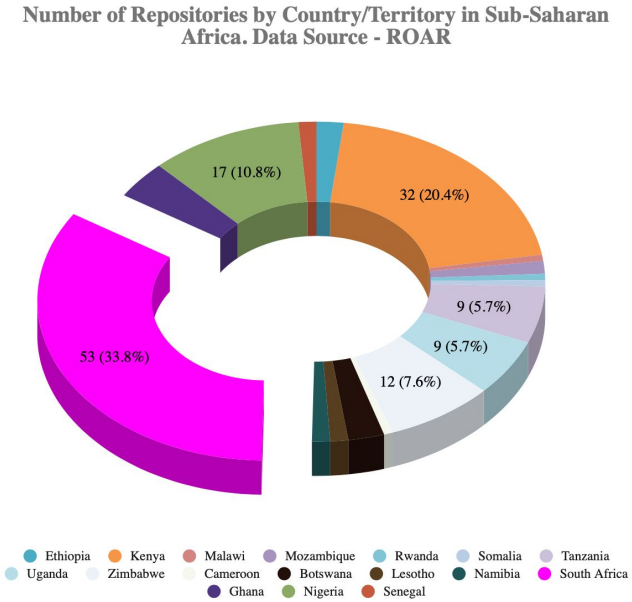
Engaging policy makers - Gwanda State University,

Librarians at AfricArxiv session with UbuntuNet

Collaborating with library communities

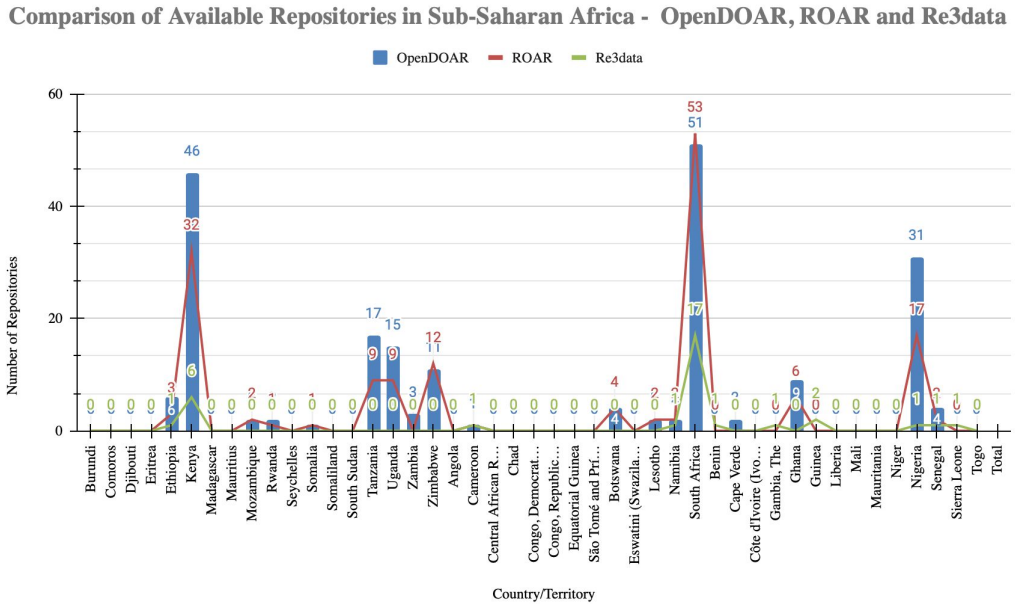
Infrastructure

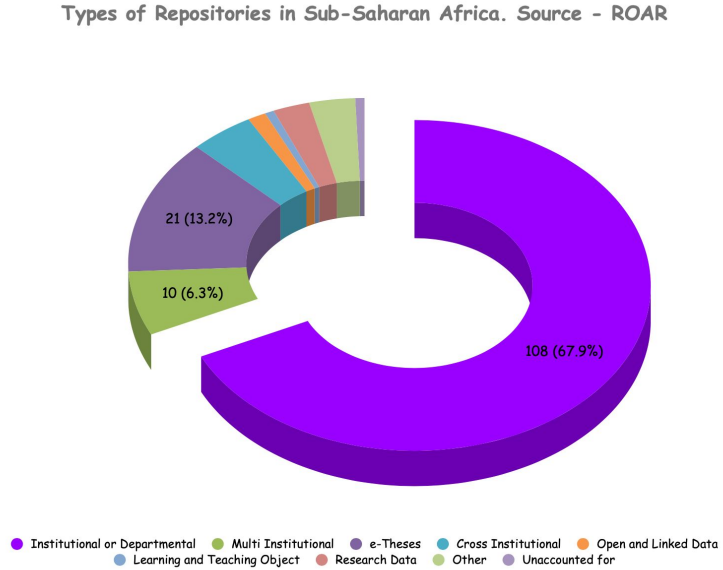
- Analyze the infrastructure landscape and use of different platforms in the region.
- Collaborate with repository platforms/SPs on providing integrations.



Infrastructure

Infrastructure Landscape Analysis





The Global Access Fund

- It aims to enable organizations worldwide to make their research outputs discoverable.
- Non profit orgs in Africa, Asia, Middle East and Latin America can apply.
- Funding categories:
  - **Outreach activities**: up to US$10,000
  - **Infrastructure development**: up to US$20,000
  - **Demonstrators**: up to US$50,000
- Applications open Sep 1 - Oct 15
- Successful applicants to be announced in Dec 2023
- Projects will run throughout 2024

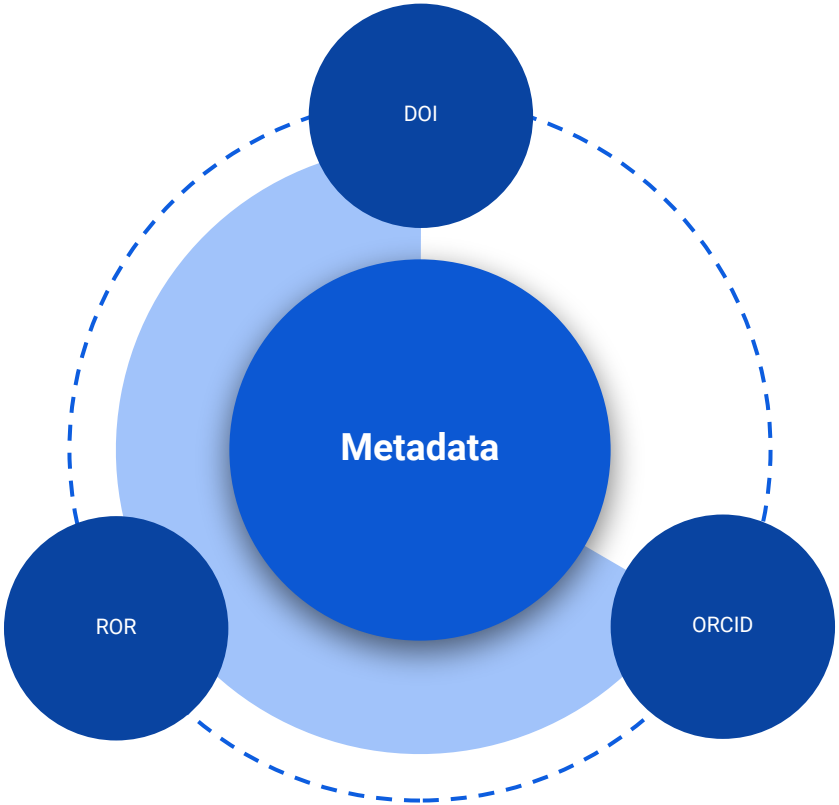
Watch out for GAF 2024

https://datacite.org/global-access-fund.html

DataCite

Connecting Research PIDs and Metadata

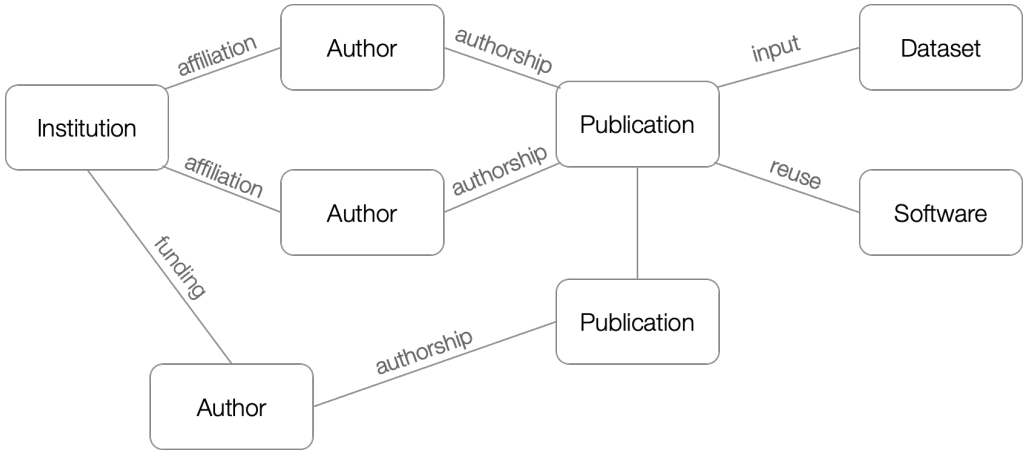
Find and connect
Metadata

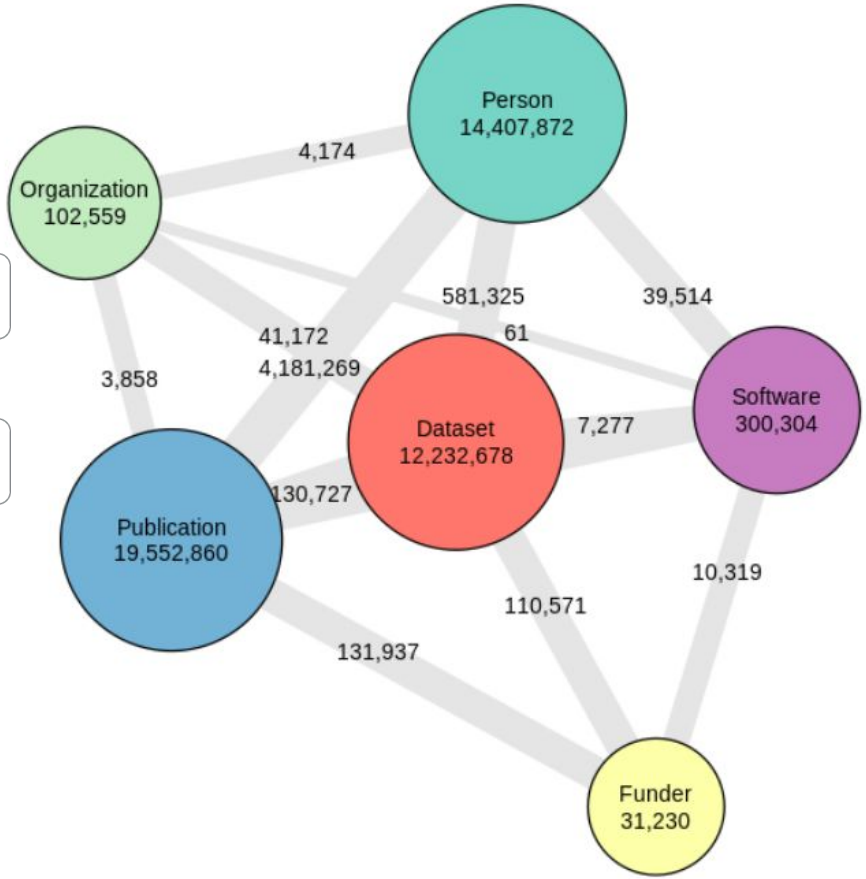


Research Impact

The PID Graph

Number of nodes and connections
(August 2022)





DataCite Metadata Schema

Metadata for DataCite DOIs

Metadata provides information about the relevant resource

When a member registers a DOI, they always register metadata about the resource

This helps with discoverability, connectedness and FAIRness

DataCite Metadata Schema

DataCite

What is the DataCite Metadata Schema?

- A list of core metadata properties chosen for an accurate and consistent identification of a resource for citation and retrieval purposes, along with recommended use instructions.
- The schema provides standardization which allows users to search across metadata, and thereby increases interoperability.
- DataCite metadata schema can be used for different resource types as shown in the next page. It work for all domains/thematic research output areas.

Metadata schema

DataCite

Add your DOI metadata following DataCite Metadata Schema

Mandatory	Recommended	Optional
Identifier	Subject	Language
Creator	Contributor	AlternateIdentifier
Title	Date	Size
Publisher	RelatedIdentifier	Format
PublicationYear	Description	Version
ResourceType	GeoLocation	Rights
		FundingReference
		RelatedItem

DataCite DOIs are suitable for a wide range of research outputs

Resource types

Audiovisual
Book
BookChapter
Collection
ComputationalNotebook
ConferencePaper
ConferenceProceeding
DataPaper
Dataset
Dissertation
Event
Image
InteractiveResource
Journal
JournalArticle
Model
OutputManagementPlan
PeerReview
PhysicalObject
Preprint
Report
Service
Software
Sound
Standard
Text
Workflow
Other

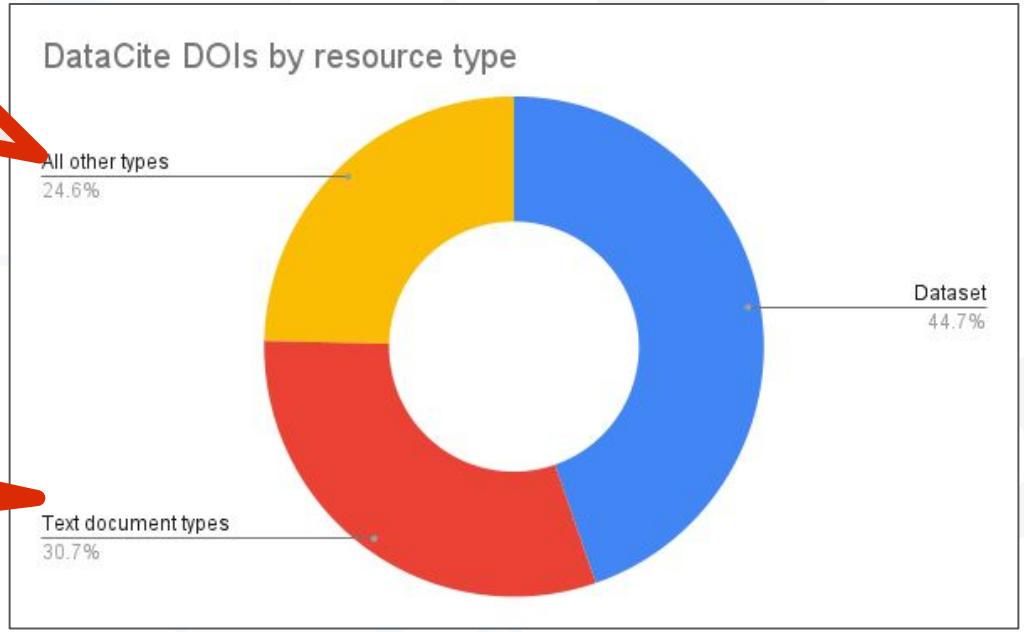
DataCite Metadata Working Group. (2021). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs. Version 4.4. DataCite e.V. https://doi.org/10.14454/3w3z-sa822

DataCite Services

Register DOI for...

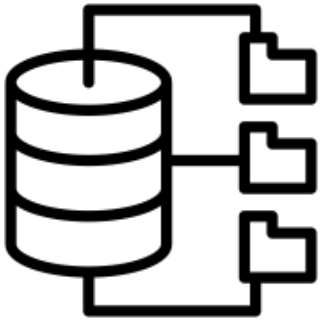
Audiovisual, Collection, ComputationalNotebook, **Dataset**, **Event,** Image, InteractiveResource, Model, **OutputManagementPlan**, PeerReview, **PhysicalObject**, Service, Software, Sound, Workflow, **Instruments**, **DMP,**Other

Book, BookChapter, ConferencePaper, ConferenceProceeding, DataPaper, Dissertation, Journal, JournalArticle, Preprint, Report, Standard, Text



Our services

- Centralized tool, Fabrica, to manage DOI activities
- Create and manage DOIs via API or webform
- Metadata Schema designed for non-traditional literature
- Search Index for all public DataCite DOIs
- Link checking service to check the health of URLs
- Create downstream impact
    - DataCite DOIs available in Google Dataset Search
    - OAI-PMH metadata made available via 3rd party indexing services
- Integration with ORCID
- Backend integration with Crossref
- APIs for various automations, reporting and dashboard



DataCite Commons

DataCite Commons

Provides free access to metrics and research output impact measurement.

Free metric platform to check:

- Works impact
- Institutions' impact
- Number of citations,
- Number of views,
- Number of downloads

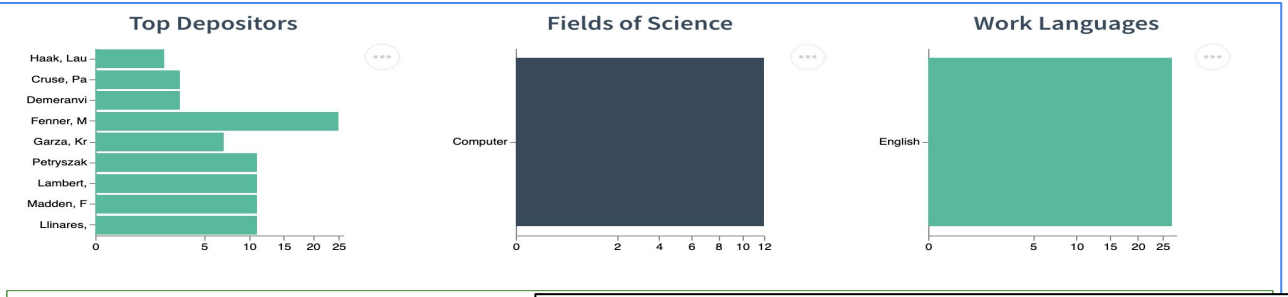
Metrics for:

- Individuals/People
- Organizations
- Works
- Repositories

Supports recognition

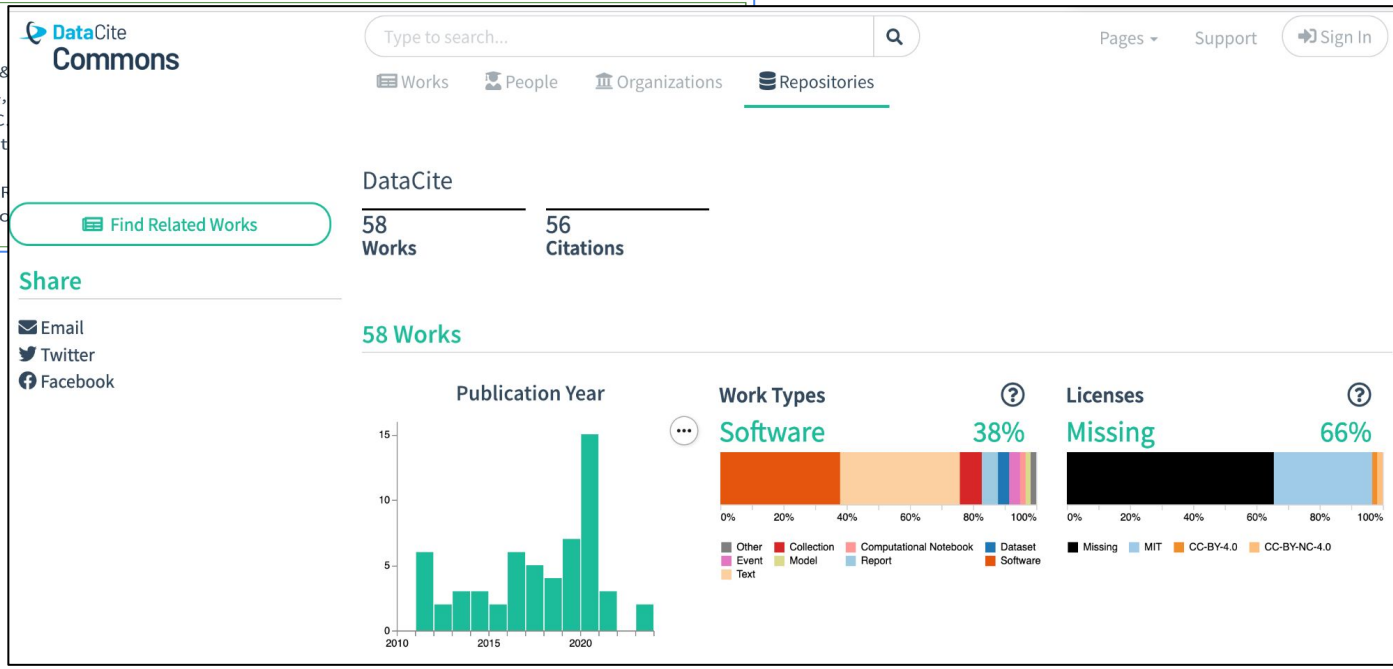
DataCite Commons

DataCite



More information about research data management

1. Borghi, J., Abrams, S., Lowenberg, D., Simms, S., &
   for Researchers. Research Ideas and Outcomes, 4,
2. Goodman, A., Pepe, A., Blocker, A. W., Borgman, C
   Care and Feeding of Scientific Data. PLoS Comput
   /journal.pcbi.1003542
3. Pampel, H., Vierkant, P., Scholze, F., Bertelmann, F
   Repositories Visible: The re3data.org Registry. PLo



Find a Journal Article

DataCite

DataCite Commons

comparative analysis of the S-locus and nuclear SSR

Pages Support Sign In

Works People Organizations

2 Works

Publication Year

- 2012 2

Work Type

- Dataset 1
- Text 1

License

- CC0-1.0 1

Language

- English 1

Registration Agency

- Crossref 1
- DataCite 1

Data from: Impact of negative frequency-dependent selection on mating pattern and genetic structure: a comparative analysis of the S-locus and nuclear SSR loci in Prunus lannesiana var. speciosa

Kato Shuri, Teruyoshi Nagamitsu, Hiroyoshi Iwata, Yoshihiko Tsumura, Yuzuru Mukai, K Michiharu, K Saika & K Junko
Version 1 of Dataset published 2012 in DRYAD

Mating processes of local demes and spatial genetic structure of island populations at the self-incompatibility (S-) locus under negative frequency-dependent selection (NFDS) were evaluated in Prunus lannesiana var. speciosa in comparison with nuclear simple sequence repeat (SSR) loci that seemed to be evolutionarily neutral. Our observations of local mating patterns indicated that male-female pair fecundity was influenced by not only self-incompatibility, but also various factors such as kinship, pollen production and flowering synchrony. In spite of the mating bias caused by these factors, the NFDS effect on changes in allele frequencies from potential mates to mating pollen was detected at the S-locus but not at the SSR loci although the changes from adult to juvenile cohorts were not apparent at any loci. Genetic differentiation and isolation-by-distance over various spatial scales were smaller at the S-locus than at the SSR loci, as expected under the NFDS. All ele sharing distributions among the populations also had a unimodal pattern at the S-locus, indicating the NFDS effect except for alleles unique to individual populations probably due to isolation among islands, although this pattern was not exhibited by the SSR loci. Our results suggest that the NFDS at the S-locus has an impact on both the mating patterns and the genetic structure in the P. lannesiana populations studied.

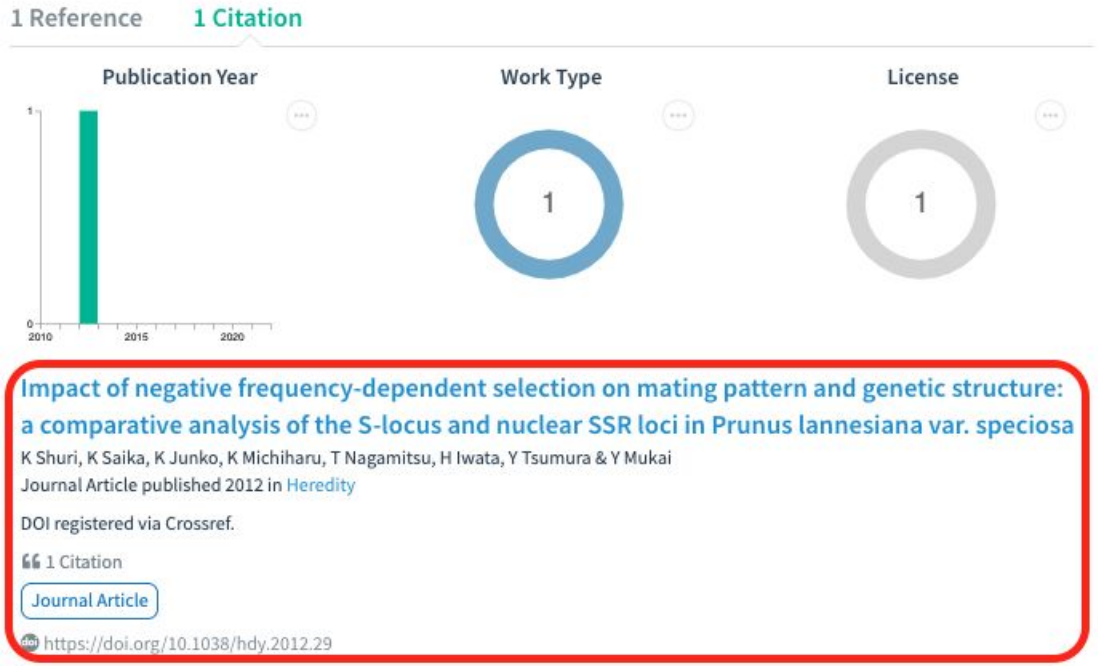
DOI registered April 17, 2012 via DataCite.

1 Citation 103 Views 16 Downloads

Dataset English

https://doi.org/10.5061/dryad.7c425

Bring citations to the surface



Support recognition

DataCite

Data from: Impact of negative frequency-dependent selection on mating pattern and genetic structure: a comparative analysis of the S-locus and nuclear SSR loci in Prunus lannesiana var. speciosa

https://doi.org/10.5061/dryad.7c425

1 Citation · 118 Views · 16 Downloads

Add to ORCID Record

Download Metadata

Description | Creators | Registration

Kato Shuri	Forestry and Forest Products Research Institute
Teruyoshi Nagamitsu	Forestry and Forest Products Research Institute
Hiroyoshi Iwata	University of Tokyo
Yoshihiko Tsumura	Forestry and Forest Products Research Institute
Yuzuru Mukai	Gifu Univ
K Michiharu	Kyoto Un
K Saika	Tokyo In
K Junko	Gunma U

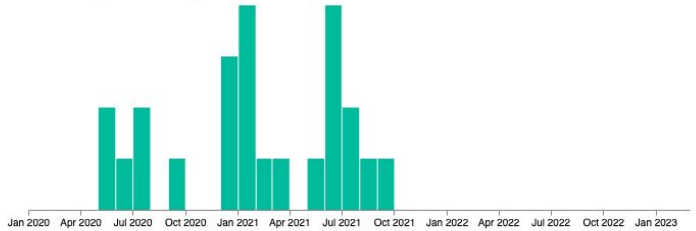
Cite as

Shuri, K., Nagamitsu, T., Iwata, H., Tsumura, Y., Mukai, Y., Michiharu, K., Saika, K., & Junko, K. (2012). *Data from: Impact of negative frequency-dependent selection on mating pattern and genetic structure: a comparative analysis of the S-locus and nuclear SSR loci in Prunus lannesiana var. speciosa* (Version 1) [Data set]. Dryad. https://doi.org/10.5061/DRYAD.7C425

APA

118 Views | 16 Downloads

118 views reported since publication in 2012.



Organizations – citations and use

World Agroforestry Centre https://ror.org/01kmz4383

2,915	29	1,116	262
Works	**Citations**	**Views**	**Downloads**

Founded 1978

Links

Homepage
Wikipedia
Twitter

Other Identifiers

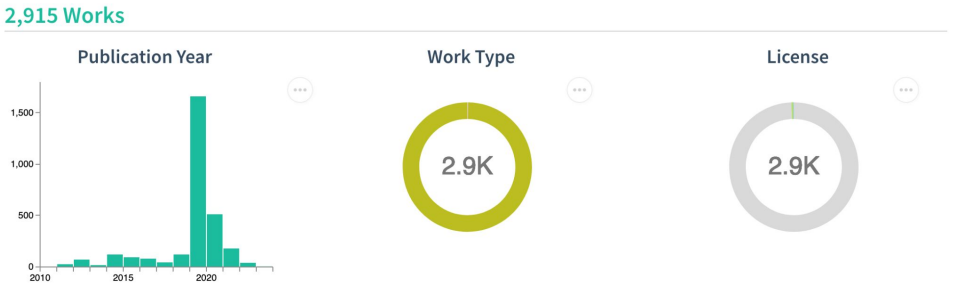
GRID grid.435643.3
Crossref Funder ID 10.13039/501100015769
ISNI 0000000099721350
Wikidata Q1362380

Geolocation

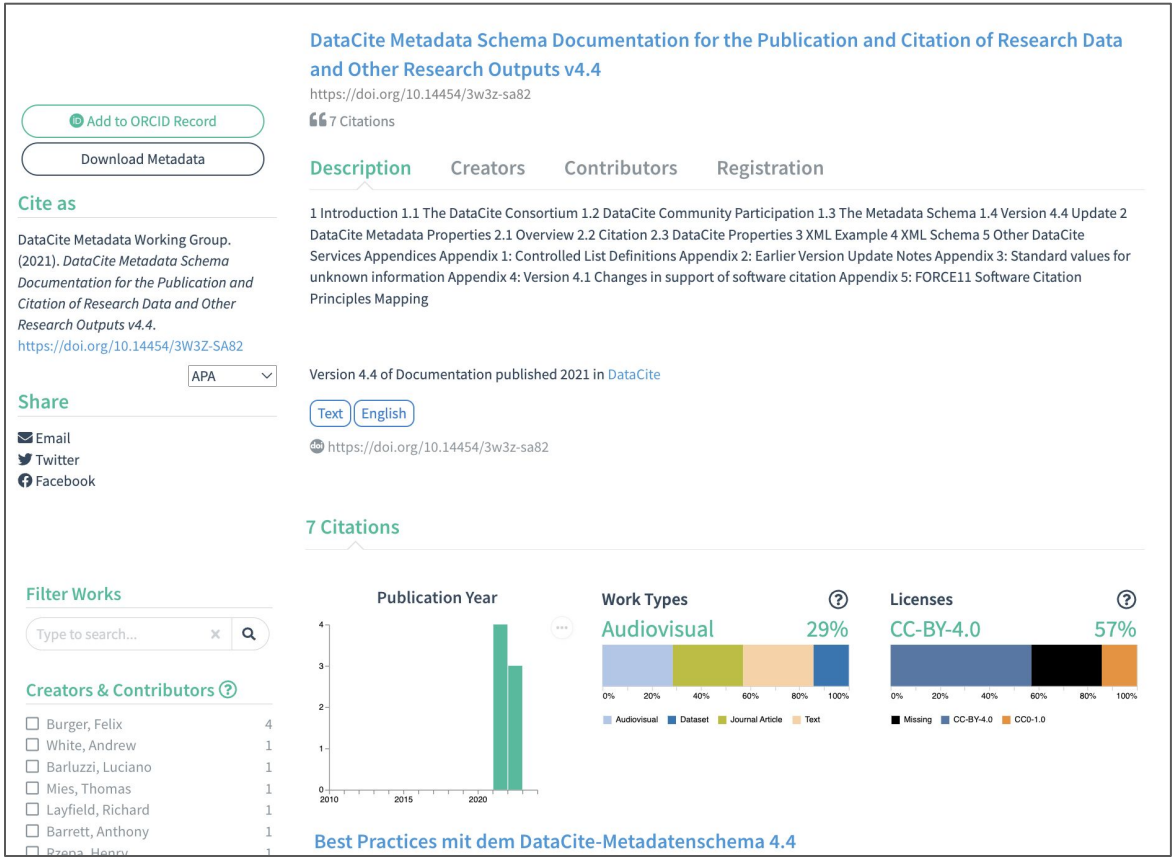
1° 14' 11.67" S36° 49' 08.4216" W

Kenya | Nonprofit | DataCite Consortium Organization

https://ror.org/01kmz4383

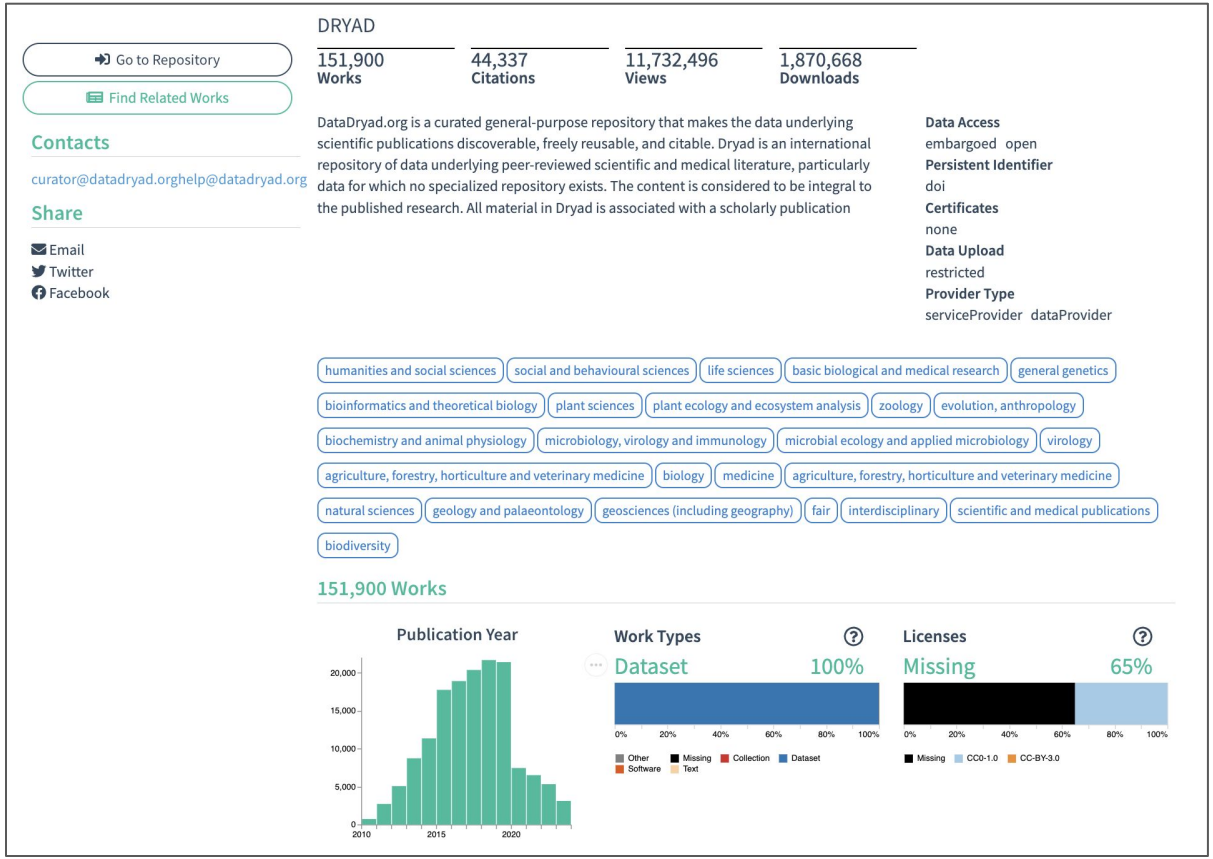


DataCite Commons: Work



https://commons.datacite.org/doi.org/10.14454/3w3z-sa82

DataCite Commons: Repository



https://commons.datacite.org/repositories/nxrc8v

Trainings and communities

DataCite Trainings and Communities

DataCite

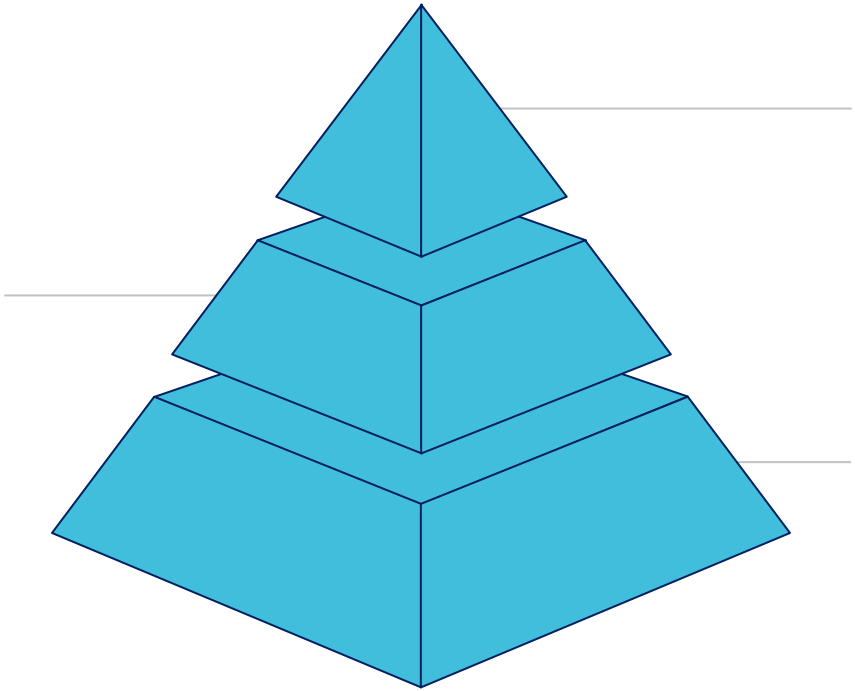
Training and Resources

- We organize regular training sessions exclusive for members
- Our past webinars and training are on our YouTube channel
- We collaborate and support members with unique needs
- We organize Open Hours to foster collaborations among our members
- We have working groups to endears new knowledge and promote best practices.
- We dedicate specialized staff for each region to provide direct engagements for our members in each region

DataCite

Membership,

Membership Benefits



Adopt and implement best practice

- Simple interfaces and services, support do documentation and dedicated staff.
- Best practice documentation.
- Community coordination, full of passionate people who share experiences and support best practice adoption.
- Continued evolution of our metadata schema.

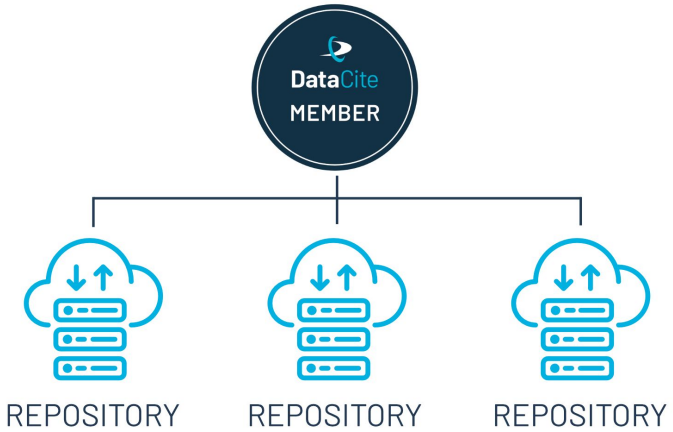
Track the influence of research with tools and services

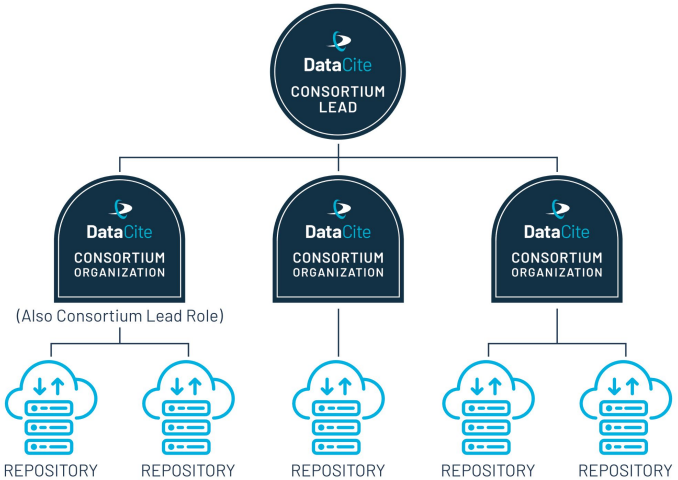
- Dashboards and analytics.
- Harvesting services.
- Graph APIs and relational metadata discovery.

Register DataCite DOIs and metadata to improve the discoverability and reuse of research outputs and resources

- DOI metadata registration and maintenance
- Content negotiation
- Link checking
- Public APIs for harvesting by third parties
- Interoperable schema

Membership Types





Feel free to contact
Bosun Obileye

bosun.obileye@datacite.org
Regional Engagement Specialist - Africa (DataCite)

DataCite

CONNECTING RESEARCH,
ADVANCING KNOWLEDGE

info@datacite.org

pidforum.org

datacite.org
blog.datacite.org

support.datacite.org
support@datacite.org

@datacite

DataCite

@datacite