The Power of PIDs: using persistent identifiers to enable connections throughout the research lifecycle

VIVO Conference June 23rd, 2021

Carly Robinson, PhD
OSTI Assistant Director
Information Products and Services



Persistent Identifiers

What is a Persistent Identifier (PID)?

- A digital identifier that is globally unique, persistent, machine resolvable, has an associated metadata schema, identifies an entity (e.g. individual researcher, publication, awards, digital research output) and is frequently used to disambiguate between entities.
- The identifier is a string of numbers, letters, and/or symbols assigned to the digital object.



What are the benefits of assigning/using PIDs?

- PIDs enable research to be more open and FAIR metadata associated with PIDs is publicly available.
- PIDs are stable, persistent links that allow for metadata to be updated as needed.
- PIDs allow organizations to track awards, research outputs, and researchers see return on investment and impact of research.
- By linking PIDs you can create connections throughout the research lifecycle connecting people to awards to organizations to research outputs to peer review.

Types of PIDs

Digital Object Identifiers (DOIs)

- Based on ISO 26324 Digital Object Identifier System and implemented on the Handle System
- International DOI Foundation management body, with Registration Agencies (RAs)
- Crossref and DataCite are RAs

Crossref

- Membership organization publishers, funders
- DOIs primarily for text-based research outputs journal articles, conference papers, technical reports
- DOIs for awards/grants/contracts

Crossref

DataCite

- Membership organization data repositories, universities, funders
- DOIs primarily for data and software



ORCID iDs



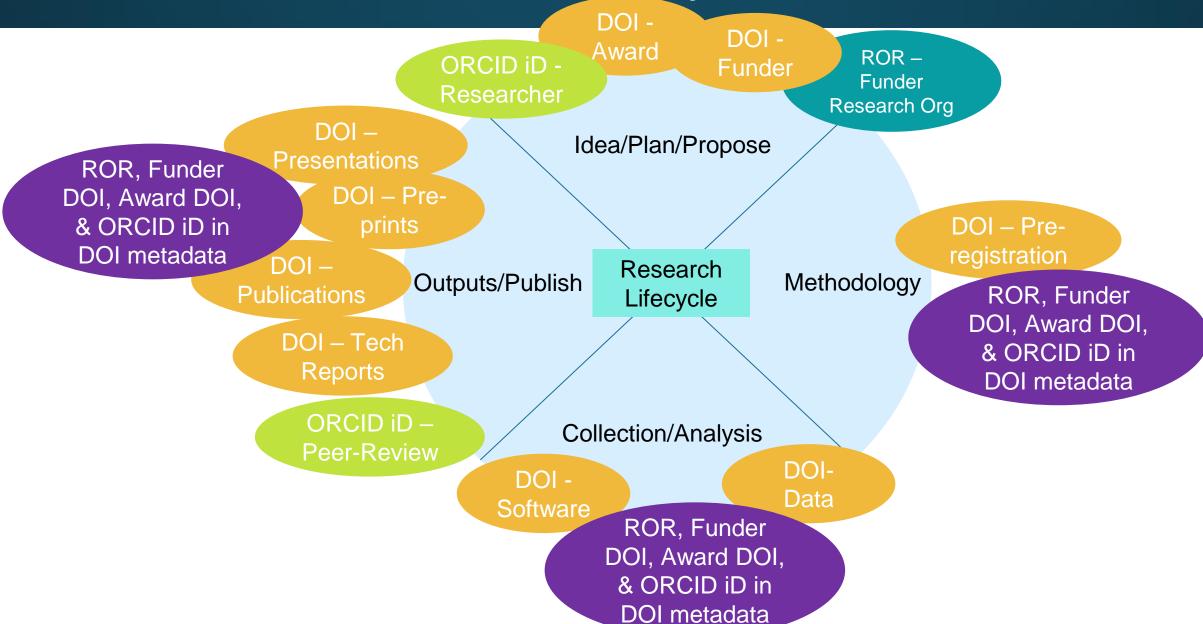
- PIDs for people/researchers
- Member organization universities, funders, government organizations

Organization Identifiers

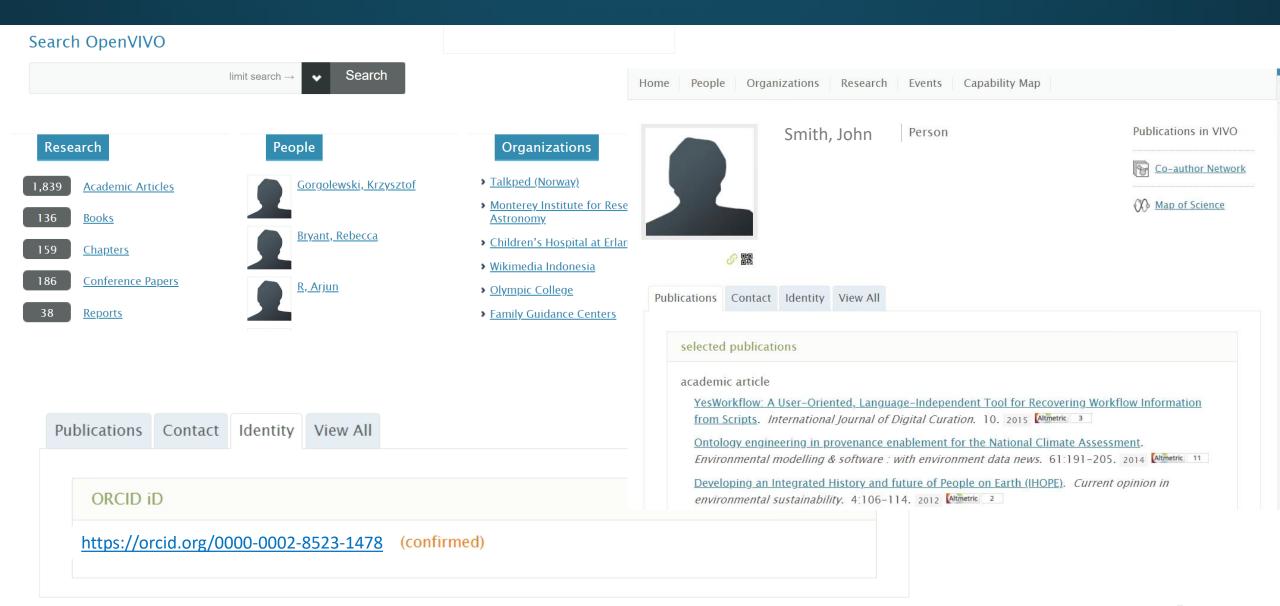
- Research Organization Registry (ROR) IDs –
 PIDs for organizations participating in research
 - Open community and infrastructure, nonprofit
- GRID IDs very broad organization ID
 - Open data, but for-profit organization
- DOIs for funding organizations Open Funder Registry



PIDs in the Research Lifecycle



Research Information Systems and PIDs



DOE Invests ~ \$12B per year in R&D



- Advanced Research Projects Agency Energy (ARPA-E)
- Office of Electricity Delivery & Energy Reliability
- Office of Energy Efficiency & Renewable Energy
- Office of Environmental Management
- Office of Fossil Energy
- Office of Legacy Management
- Office of Nuclear Energy
- Office of Science
- Office of Environment, Health, Safety & Security

NATIONAL LABS

Ames

Argonne

Brookhaven

Fermi

Idaho

Los Alamos

Lawrence Berkeley

Lawrence Livermore

NETL

NREL

Oak Ridge

Pacific Northwest

Princeton

SLAC

Sandia

Savannah River

Thomas Jefferson

GRANTEES TECHNOLOGY CENTERS SITES

R&D Results



- Journal articles/accepted manuscripts
- Technical reports
- Conference papers
- Theses/dissertations
- Scientific and technical software
- Data objects
- Patents
- Workshop reports
- Videos

≈ 50,000 R&D findings/results per year

OSTI Mission and Priorities

Core Functions: OSTI collects, preserves, and disseminates DOE-funded R&D results.

Mission: Advance science and sustain technological creativity by making R&D findings available and useful to Department of Energy researchers and the public.

Strategic Priority: Provide our community with persistent identifier (PID) services.



DOE OSTI PID Services

Service Name	Research Object	Service Partner
PIDs for Research Results		
E-Link (research output ingest system)	Technical/Workshop Reports Conference Posters Presentations	Crossref
DOE Data ID Service	Data	DataCite
Interagency Data ID Service (IAD)	Data/Research Outputs	DataCite
DOE CODE	Software	DataCite
PIDs for Awards		
Award DOI Service	Awards	Crossref Grant ID
PIDs for People		
US Government ORCID Consortium	Researchers	ORCID
PIDs for Organizations		
Organization Authority	Research/Funding Organizations	ROR
Open Funder Registry	Funding Organizations	Crossref/Elsevier

Data ID Services





Award DOI Service

US Government ORCID Consortium

PIDs for Research Results – DOIs

Data ID Services — Data DOIs

DOE Data ID Service

- Labs, facilities, and data repositories provide data records (with associated metadata) to OSTI via E-Link.
- E-Link is the corporate system for collecting DOE-funded research results, developed and maintained by OSTI.
- When data record is submitted, OSTI automatically assigns a DataCite DOI (unless one is already assigned).

DOE Data ID Service

Interagency

Data ID Service

Interagency Data ID Service (IAD)

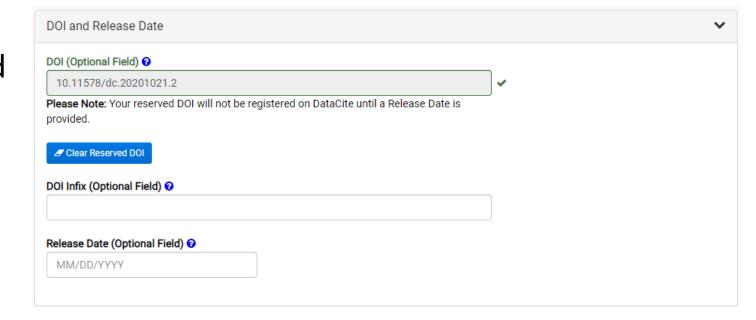
- Assign DataCite DOIs to other agencies' research outputs, based on metadata passed from OSTI to DataCite.
- Provide service on a cost recovery basis.
- Currently working with 5 agencies DOT, USDA, NASA (PDS), NIH (NIMH and BRICS), and EPA.

https://www.osti.gov/data-services

DOE CODE – Software DOIs

DOE CODE is the software services platform and search tool for DOE-funded code.

- When early development code/software records submitted to DOE CODE, can optionally assign DataCite DOIs. Typically for DOI citation purposes.
- DataCite DOIs are automatically assigned to software formally announced to DOE OSTI (if one has not already been assigned).



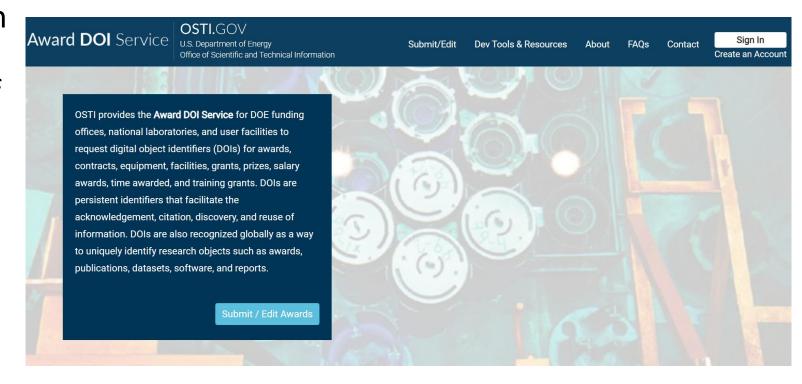
PIDs for Awards – DOIs

Award DOI Service

Service Development and Launch

https://www.osti.gov/award-doi-service/

- Crossref members can use award/grant DOI service.
- Worked with DOE user facilities to gather requirements.
- Launched the Award DOI Service pilot project September 2020.
- Piloting with facilities to assign award DOIs to the awards provided by the facility (use of the facility).
- The service can scale to support DOE funding offices and other government awards.



PIDs for People – ORCID iDs

US Government ORCID Consortium

Consortium Development and Launch

- Many DOE and US government organizations joining ORCID as direct members.
- Interest from those organizations to be an ORCID member through a consortium.
- Developed to create community, decrease costs, and providing increased services.
- US Government ORCID Consortium launched April 1st, 2020. Led by DOE OSTI.

Consortium Benefits

- Consortium members are premium ORCID members 5 API credentials, increased API functionality, custom analytics, etc.
- Consortium provides more individualized support both administrative and technical.
- Consortium members are part of the US government community of practice information sharing with organizations with similar needs and use cases.

US Government
ORCID Consortium

https://www.osti.gov/orcid-consortium/

Consortium Members

12 current consortium members















The Advanced Photon Source

Argonne Leadership Computing Facility



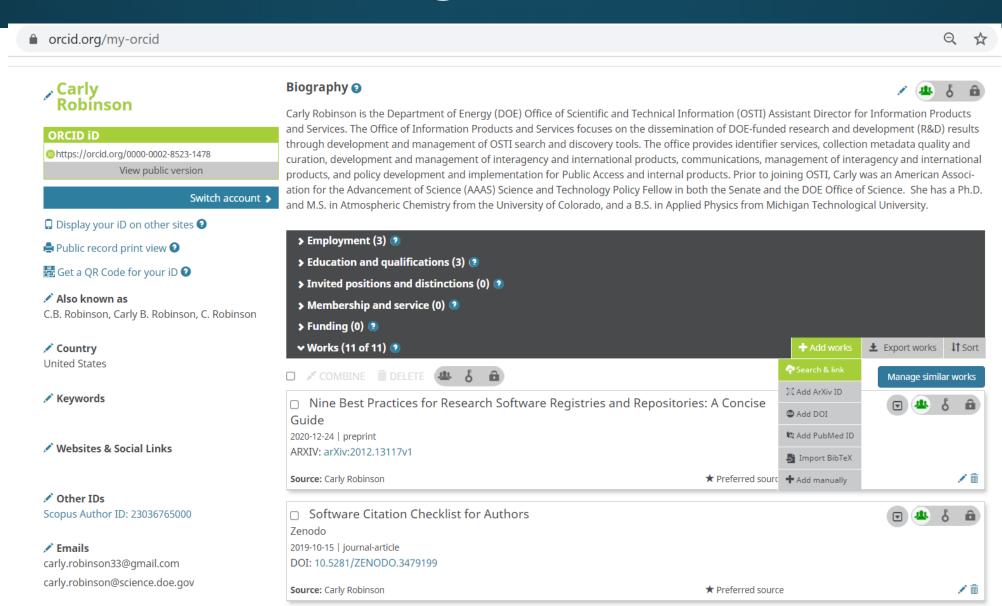
Argonne Research Library

Center for Nanoscale Materials CNM



ORCID Member Integration

^



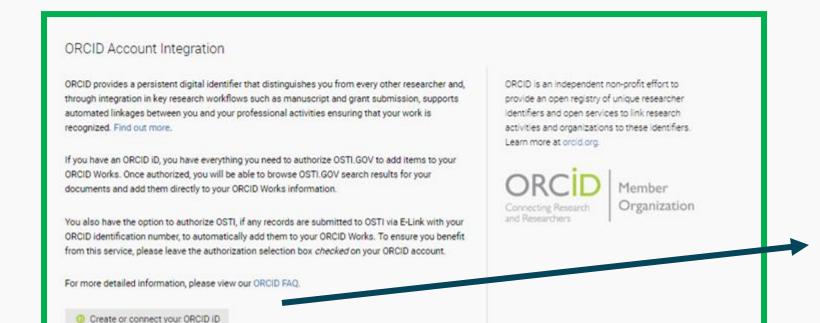
☐ Software Citation Checklist for Developers

Sign In

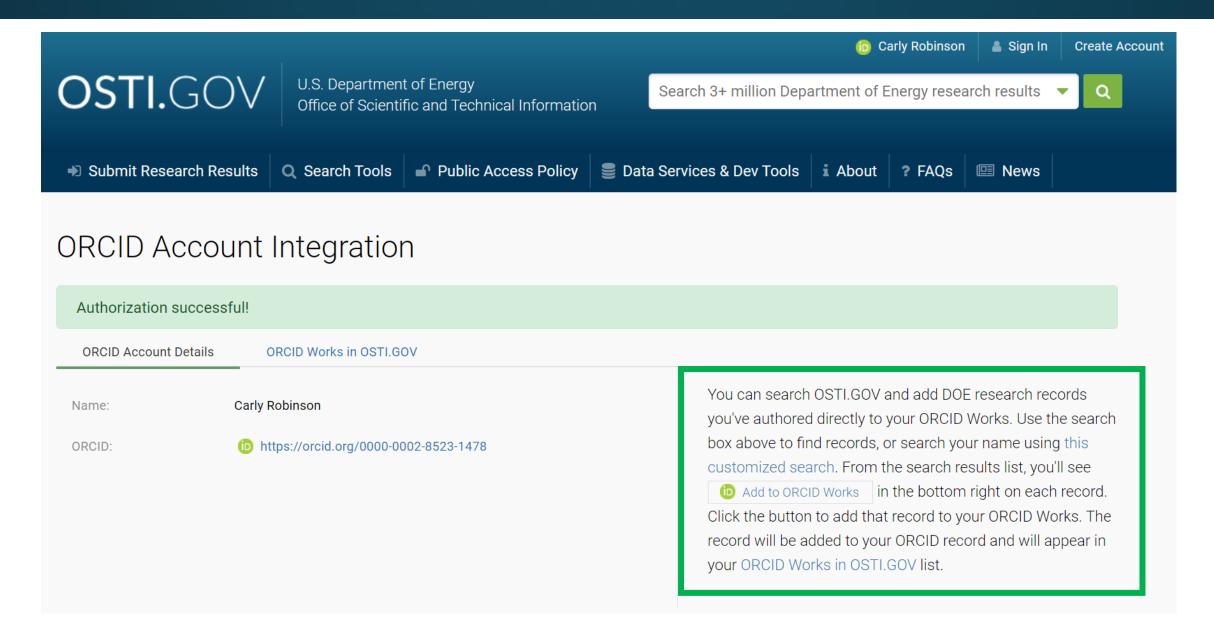
If you already have an OSTLGOV or DOE PAGES account, enter your email address and password below to sign in. OSTI has recently added new features for members of the DOE community that may be activated with some additional account information. When signed in, please visit your account management screen to ensure that your account information is complete and up-to-date. You may view saved searches, export bibliographies and create content alerts from your account management screen. Your current login credentials may be used across OSTI discovery tools.

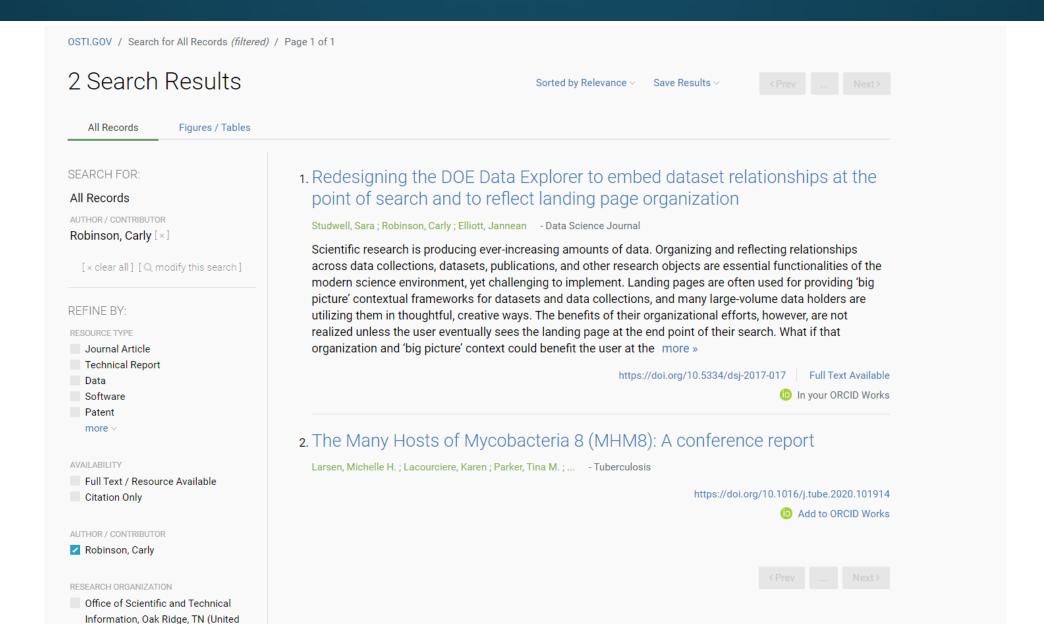


- OSTI.GOV integrates with ORCID API to allow users (researchers) to "claim" OSTI.GOV records and add to ORCID record.
- OSTI.GOV gets permission from user to push OSTI.GOV records to their ORCID record through authentication.











1. Redesigning the DOE Data Explorer to embed dataset relationships at the point of search and to reflect landing page organization

Studwell, Sara; Robinson, Carly; Elliott, Jannean - Data Science Journal

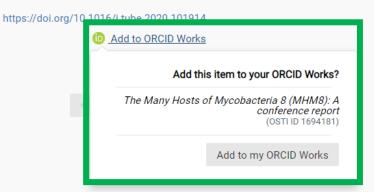
Scientific research is producing ever-increasing amounts of data. Organizing and reflecting relationships across data collections, datasets, publications, and other research objects are essential functionalities of the modern science environment, yet challenging to implement. Landing pages are often used for providing 'big picture' contextual frameworks for datasets and data collections, and many large-volume data holders are utilizing them in thoughtful, creative ways. The benefits of their organizational efforts, however, are not realized unless the user eventually sees the landing page at the end point of their search. What if that organization and 'big picture' context could benefit the user at the more 's

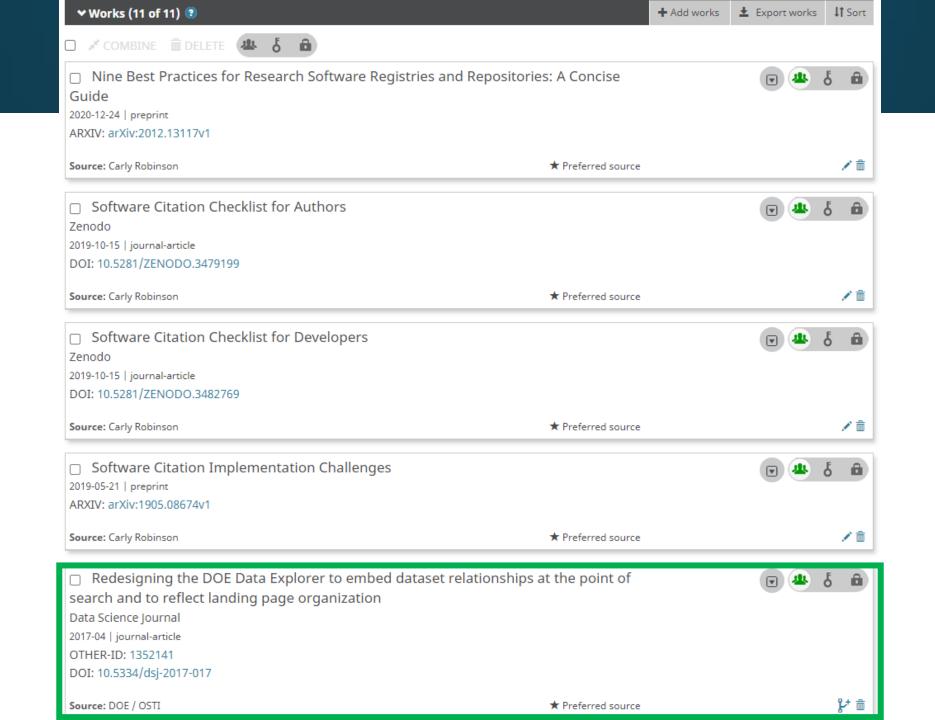
https://doi.org/10.5334/dsj-2017-017 Full Text Available

in your ORCID Works

2. The Many Hosts of Mycobacteria 8 (MHM8): A conference report

Larsen, Michelle H.; Lacourciere, Karen; Parker, Tina M.; ... - Tuberculosis

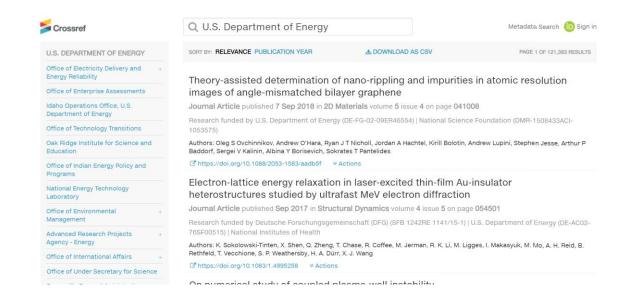




PIDs for Organizations

Open Funder Registry – DOIs for Funding Orgs

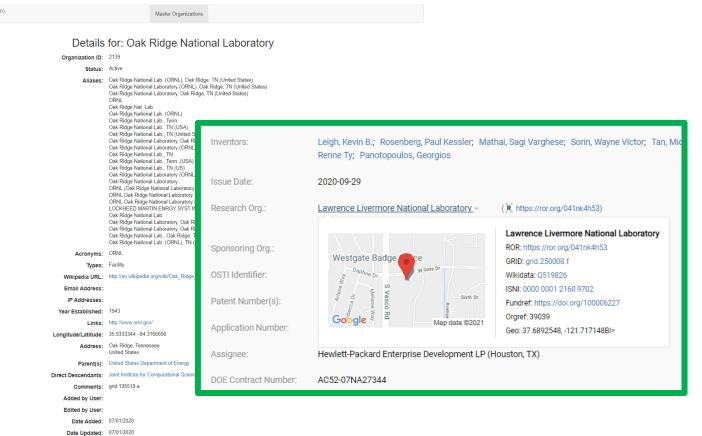
- Funder Registry maintained by Crossref and Elsevier.
- Assigns DOIs to funding organizations in hierarchy.
- OSTI maintains the DOE funding office hierarchy and provide updates as needed.
- DOIs used by publishers to identify funding organizations and implement public access requirements.
- Publishers identify funding information in journal article acknowledgements or when collected upon submission and associate the Funder DOI from the registry and include in journal article DOI metadata.



https://www.crossref.org/services/funder-registry/

Organization Authority — Org IDs

- OSTI has maintained organization authority for many years capturing sponsoring organizations, research organizations, and affiliations provided to OSTI in the records we collect.
- Developing new authority mapped to organization identifiers (ROR, GRID, Wiki).
- Developing reconciliation API for others to use to run against their own organization authorities to add identifiers.



	External ID's	
Туре	Value	Preferred?
FundRef	100006228	Yes
GeoNames City ID	4646571	Yes
GRID	grid.135519.a	Yes
ISNI	0000 0004 0446 2659	Yes
ORGRef	38147	Yes
ROR	https://ror.org/01qz5mb56	Yes
Wikidata	Q714439	Yes
Wikidata	Q6971313	No

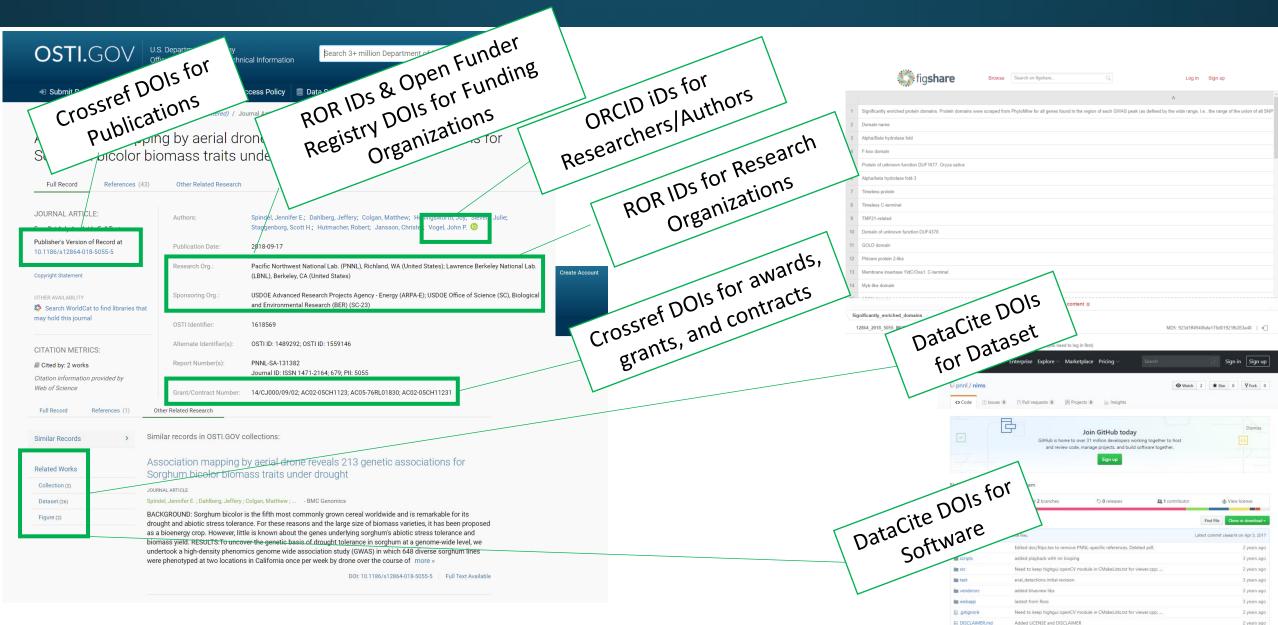
Using and Connecting PIDs

Connecting PIDs in Metadata



At DOE OSTI, we're working to create connections using PIDs throughout the research lifecycle in connected metadata to show the impact of DOE funding and tell DOE's story.

Connecting PIDs in Metadata



Visualizing PIDs

An official website of the United States government Here's how you know >

PIDs@ About DOI Services > ORCID Services > Org IDs Using PIDs News Create Account Sign In

OSTI.GOV

 $B \to \pi \ell v$ and $B_s \to K \ell v$ form factors and $|V_{ub}|$ from 2+1-flavor lattice QCD with domain-wall light quarks and relativistic heavy quarks

Persistent / Related Identifier Connections

OSTI ID: 1180799

PUBLICATION YEAR

2015

ABSTRACT

We calculate the form factors for $B \to \pi \ell V$ and $B s \to K \ell V$ decay in dynamical lattice quantum chromodynamics (QCD) using domain-wall light quarks and relativistic b-quarks. We use the (2+1)-flavor gauge-field ensembles generated by the RBC and UKQCD collaborations with the domain-wall fermion action and lwasaki gauge action.

see more

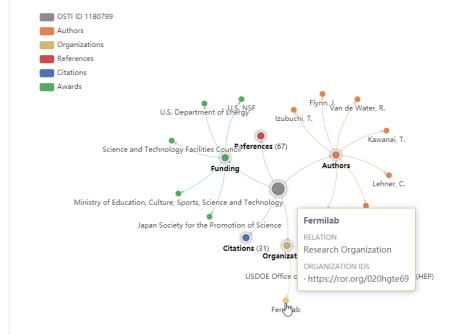
PUBLISHER'S VERSION OF RECORD

https://doi.org/10.1103/PhysRevD.91.074510

CITATION METRICS

■ Cited by:31Impact Factor:4.833Citation Impact by Journal:2.64Citation Impact by Field:2.9717% Rank by Field / Year:5.5053

Citation information provided by Web of Science



NODE DETAILS

FERMILAB

RELATION

Research Organization

http://www.fnal.gov/

Batavia, Illinois United States

41.8319435, -88.257225



ROI

https://ror.org/020hgte69

GRI

grid.417851.e

Funder Registry

100006230

ISNI

0000 0001 0675 0679

ORGRef

53301

Wikidata

Q337641

Thank you!

Carly Robinson

Carly.Robinson@science.doe.gov



www.osti.gov



pids@osti.gov



@osti.gov