DataCite

CONNECTING RESEARCH,
ADVANCING KNOWLEDGE

FAIR Facilities and Instruments in DataCite Services

Maria Gould, Director of Strategic Programs & Partnerships

22 September 2025

FAIR Facilities and Instruments Workshop

maria.gould@datacite.org

@datacite.org

About me

Current work

- *DataCite:* Director of Strategic Programs & Partnerships
- *ROR* (Research Organization Registry): Director

Prior work

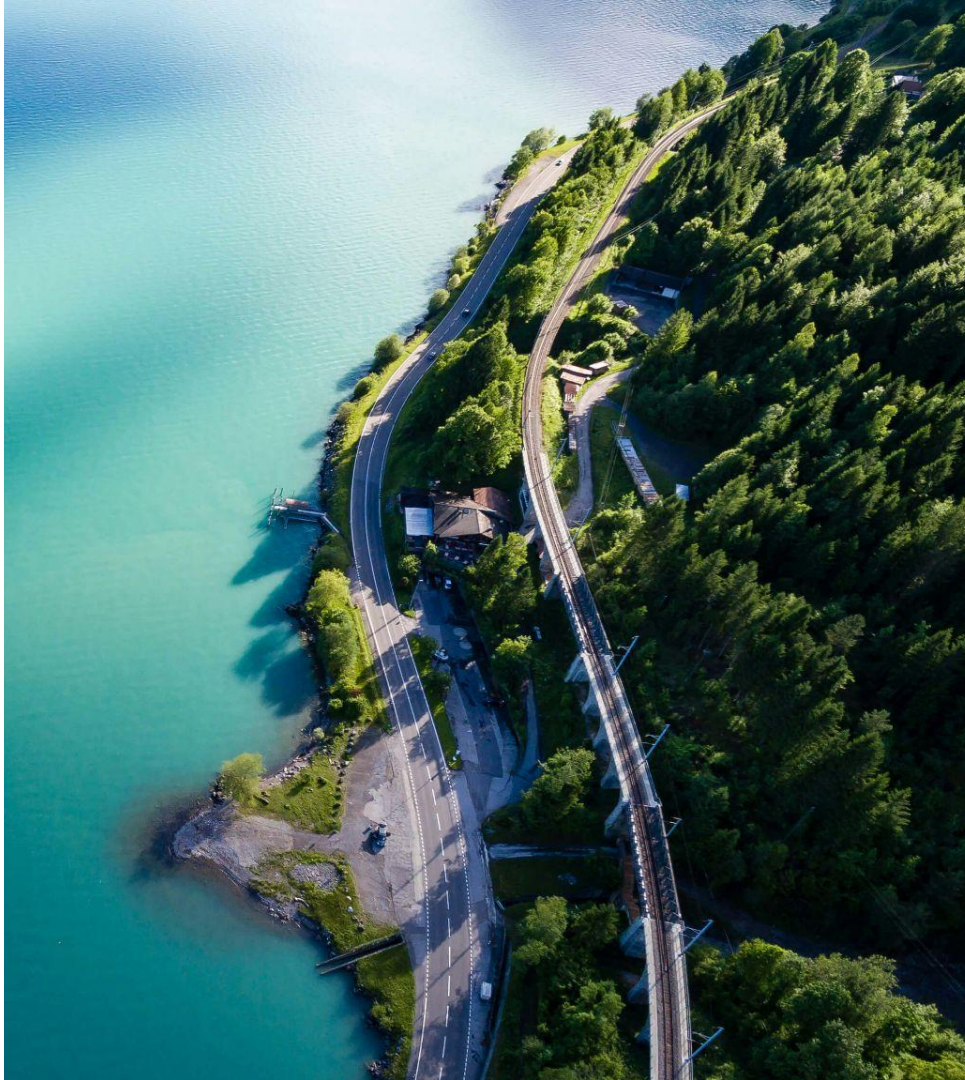
- California Digital Library / EZID
- UC Berkeley Library
- PLOS

Other relevant affiliations & activities

- Make Data Count Advisory Group
- NISO US National PID Strategy WG
- OpenAlex Advisory Board

Roadmap

- Introduction to DataCite
- Overview of facility and instrument use cases at DataCite
- Reflections on workshop recommendations
- Requests for community input and involvement



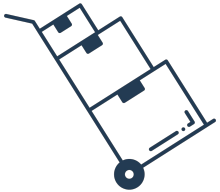
DataCite at a glance

- Global non-profit **membership organization & community of practice**
  - **Libraries, repositories, publishers, government agencies**, **research facilities** and other research organizations
- **DOI registration service** for **outputs, resources, and activities**
  - Publications, datasets, software, samples and specimens, instruments, projects, data management plans, awards, and more
- Open tools and services for **connecting, querying, discovering, analyzing, and integrating DOI and PID metadata**
  - PID graph, REST API, OAI-PMH, data dump, search and dashboard UIs
- **Standard, extensible, interoperable metadata schema**
  - Developed and maintained with user and community input, and aligned to other schemas

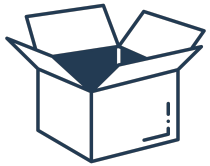
DataCite metadata:
Powering discovery

DataCite

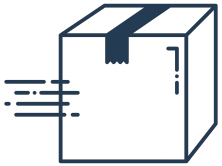
Research organizations provide metadata to DataCite and register a DOI for diverse types of outputs and activities that they produce and manage



DataCite stores the metadata and establishes connections across outputs and activities using PIDs



Metadata is made openly available through DataCite APIs, data files, and interfaces



Metadata is discovered and harvested by users and systems to track and analyze outputs or integrate in other systems

DataCite use cases and benefits

- **Create DOIs** for research outputs, resources, and activities
- Elevate **visibility and discoverability** of research
- **Index and harvest metadata** about research outputs, resources, and activities
- **Uncover connections, trends, and insights** across research outputs and activities
- **Create efficiencies** in research management, tracking, and reporting
- **Align with recommended practices and policies** for FAIR data, open science, public access, and more

DataCite key themes and takeaways

A global, dynamic community for everyone

- **Community-led non-profit open infrastructure**
- Metadata, services, and practices **always evolving alongside community needs**

An expansive and interoperable metadata store

- **More than data**
- A metadata store; **not a content repository**
- DataCite schema can **co-exist with local metadata schemes**
- Schema **designed to support citation & discovery**

One layer of a rich research landscape

- DataCite metadata **underpins many other systems**
- DataCite metadata becomes **more useful when it includes other PIDs**
- **Collaboration and coordinatio**n are key

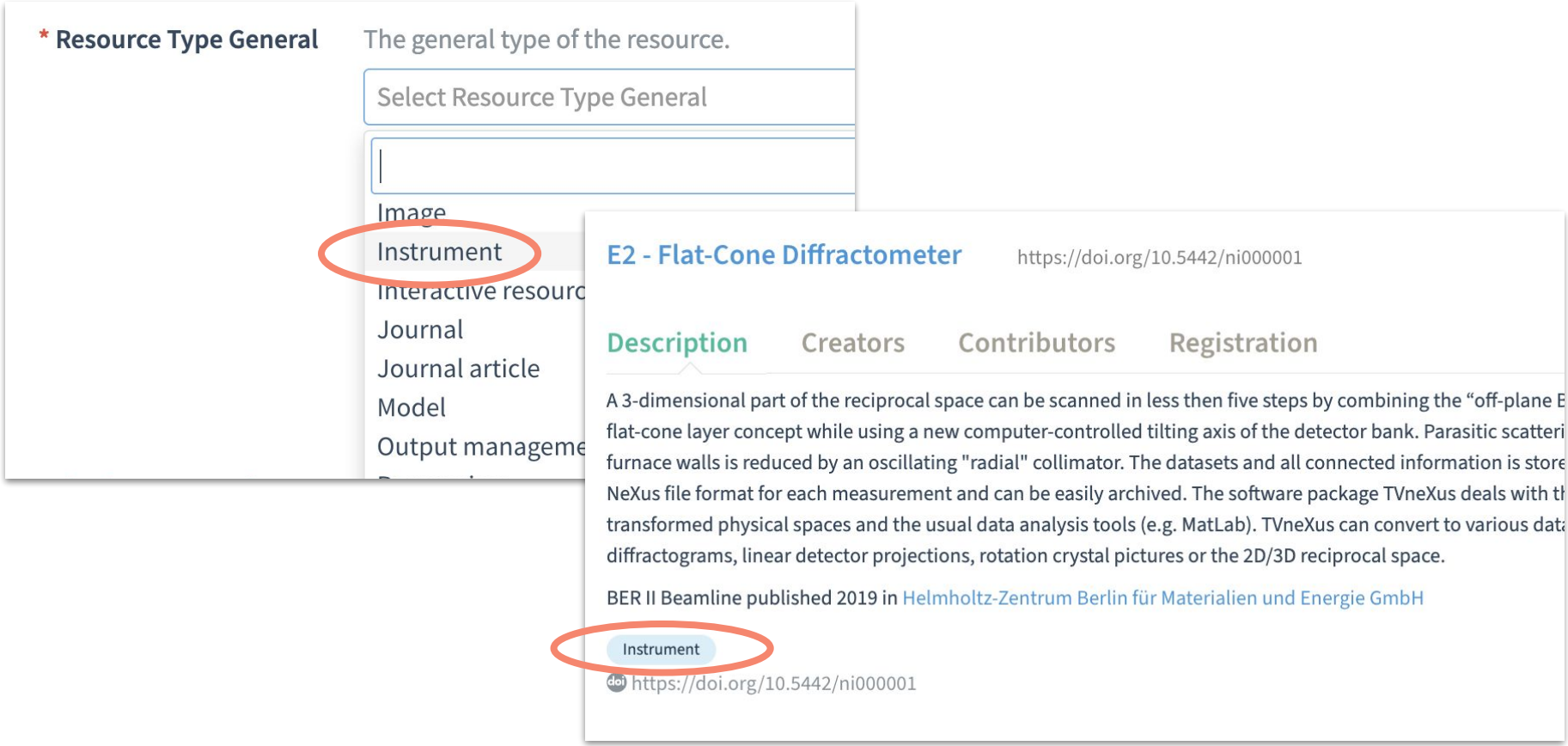
DataCite for facilities and instruments

Facilities and instruments in DataCite: multiple pathways and approaches

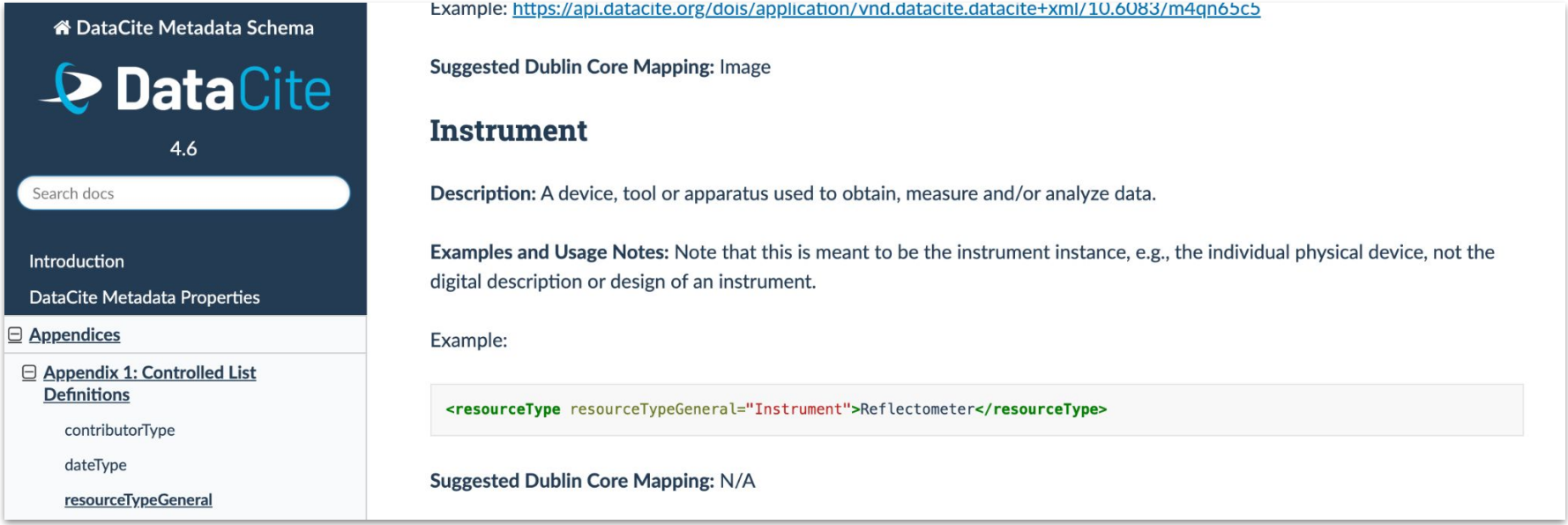
DataCite

- Resource and output **identification**
- **Attribution and relationships** via metadata schema and PID connections
  - Creator and contributor attribution
  - Institutional attribution (as creator, contributor, publisher, or funder)
  - Funder and funding acknowledgment
  - Related works and identifiers
  - Multiple relationship types
- **Metadata interoperability** via crosswalks/mappings
- **Tracking, reporting, and analysis** via APIs and discovery tools

Resource and output identification



Resource and output identification



Metadata schema: Attribution

DataCite

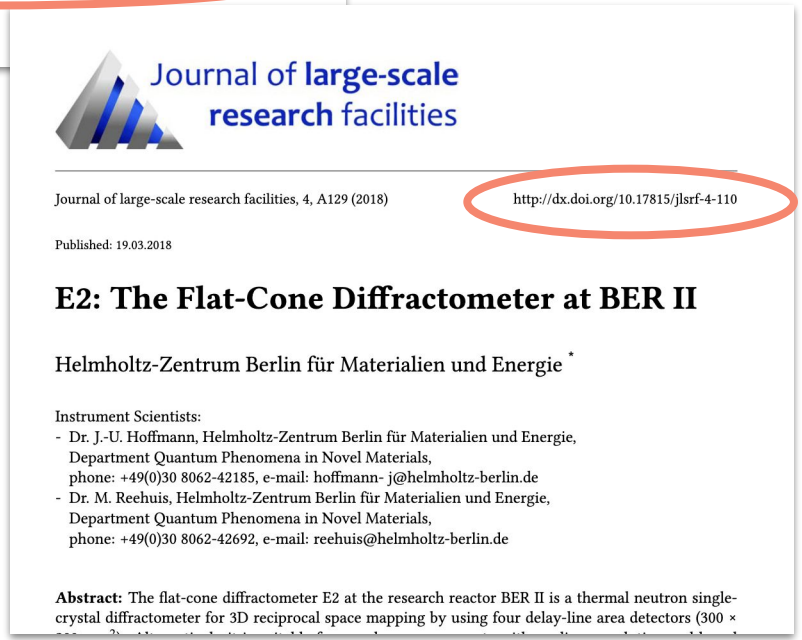

```
"creators": [
  {
    "name": "Helmholtz-Zentrum Berlin Für Materialien Und Energie",
    "nameType": "Organizational",
    "affiliation": [],
    "nameIdentifiers": [
      {
        "schemeUri": "https://ror.org",
        "nameIdentifier": "https://ror.org/02aj13c28",
        "nameIdentifierScheme": "ROR"
```


Instrument DOI metadata attributes research facility and references facility ROR ID

Metadata schema: Relationships

```
"relatedIdentifiers": [
  {
    "relationType": "IsDescribedBy",
    "relatedIdentifier": "10.17815/jlsrf-4-110",
    "relatedIdentifierType": "DOI"
```

Instrument DOI metadata references published article



Journal of large-scale research facilities

Journal of large-scale research facilities, 4, A129 (2018) http://dx.doi.org/10.17815/jlsrf-4-110

Published: 19.03.2018

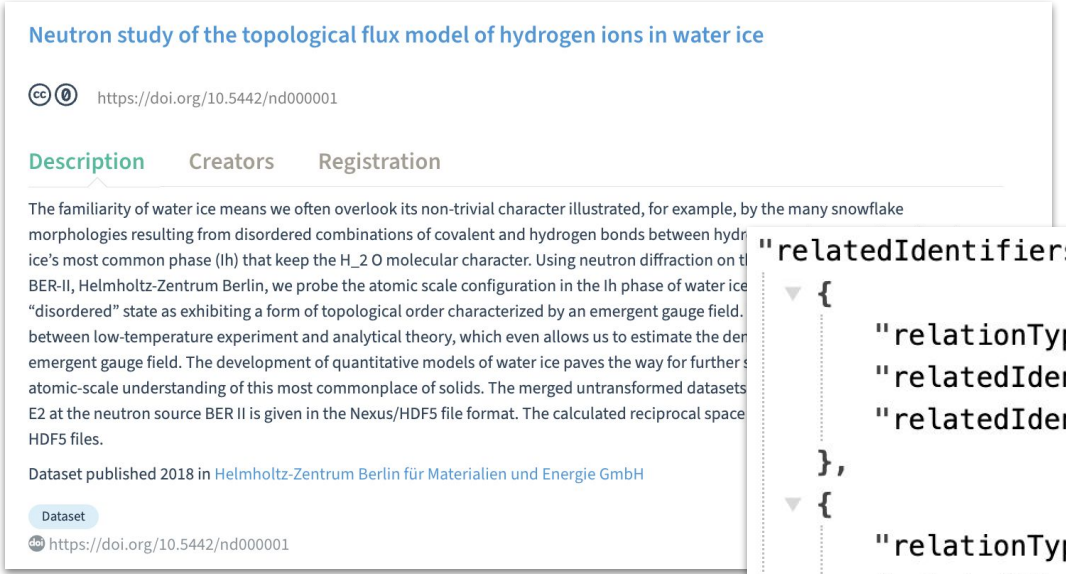
E2: The Flat-Cone Diffractometer at BER II

Helmholtz-Zentrum Berlin für Materialien und Energie *

Instrument Scientists:
- Dr. J.-U. Hoffmann, Helmholtz-Zentrum Berlin für Materialien und Energie, Department Quantum Phenomena in Novel Materials, phone: +49(0)30 8062-42185, e-mail: hoffmann- j@helmholtz-berlin.de
- Dr. M. Reehuis, Helmholtz-Zentrum Berlin für Materialien und Energie, Department Quantum Phenomena in Novel Materials, phone: +49(0)30 8062-42692, e-mail: reehuis@helmholtz-berlin.de

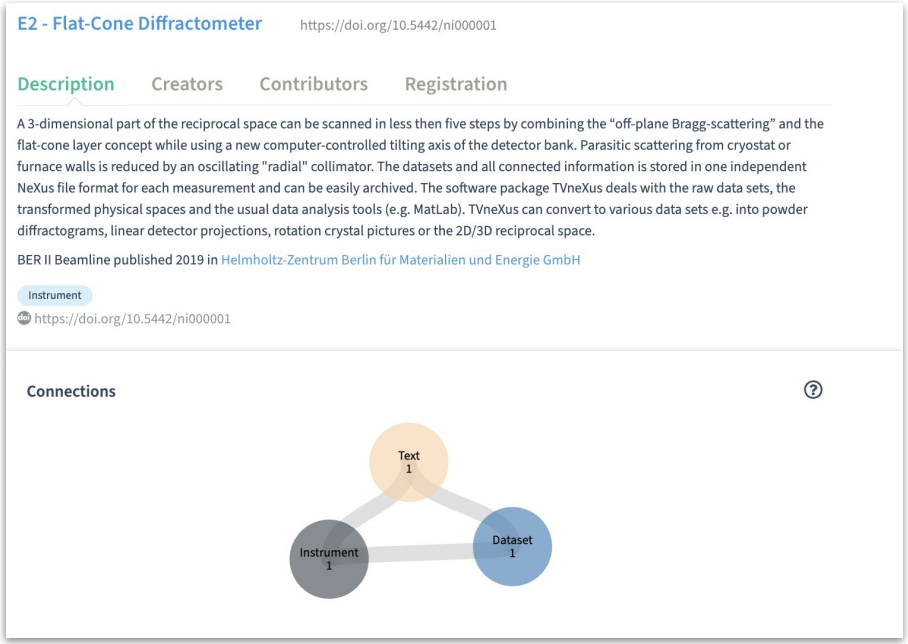
Abstract: The flat-cone diffractometer E2 at the research reactor BER II is a thermal neutron single-crystal diffractometer for 3D reciprocal space mapping by using four delay-line area detectors (300 ×

Metadata schema: Relationships



Dataset DOI metadata references instrument and published article, asserting precise relation types

Metadata schema: Relationships



Relationships exposed in DataCite Commons discovery interface

Metadata schema: Crosswalks

DataCite

DataCite Metadata Schema

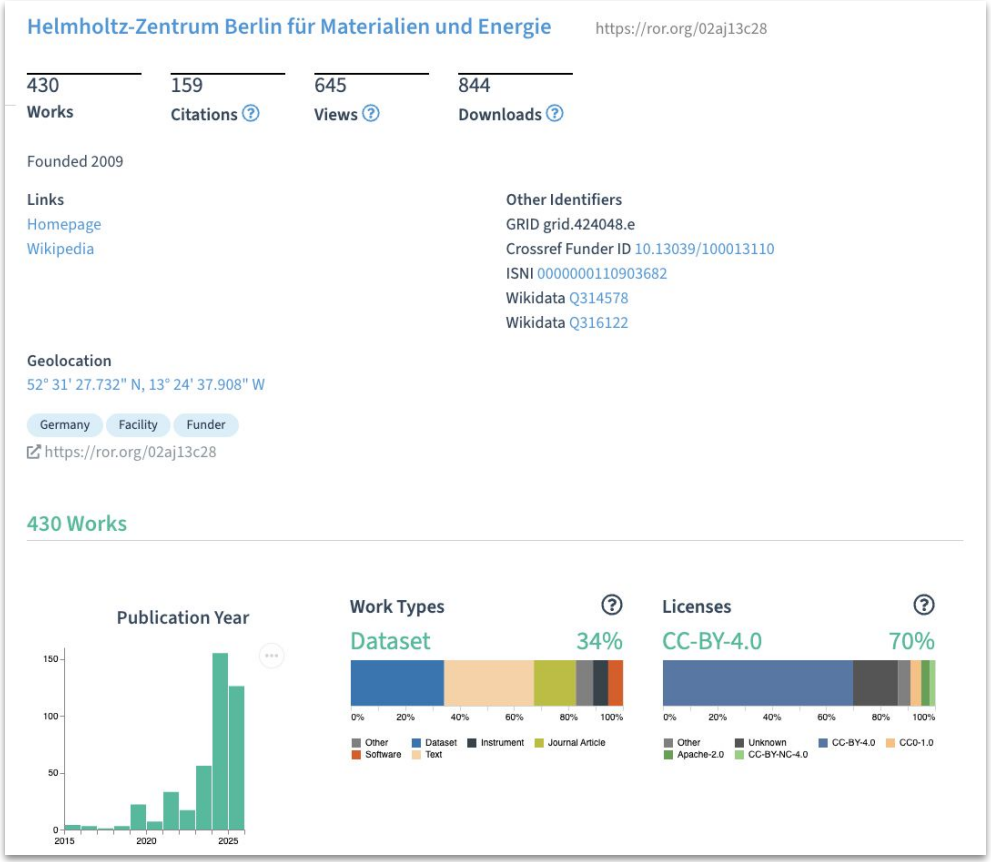
/ Mappings / PIDINST Schema Mapping

PIDINST Schema [1] Mapping

Table 7: PIDINST to DataCite Mapping

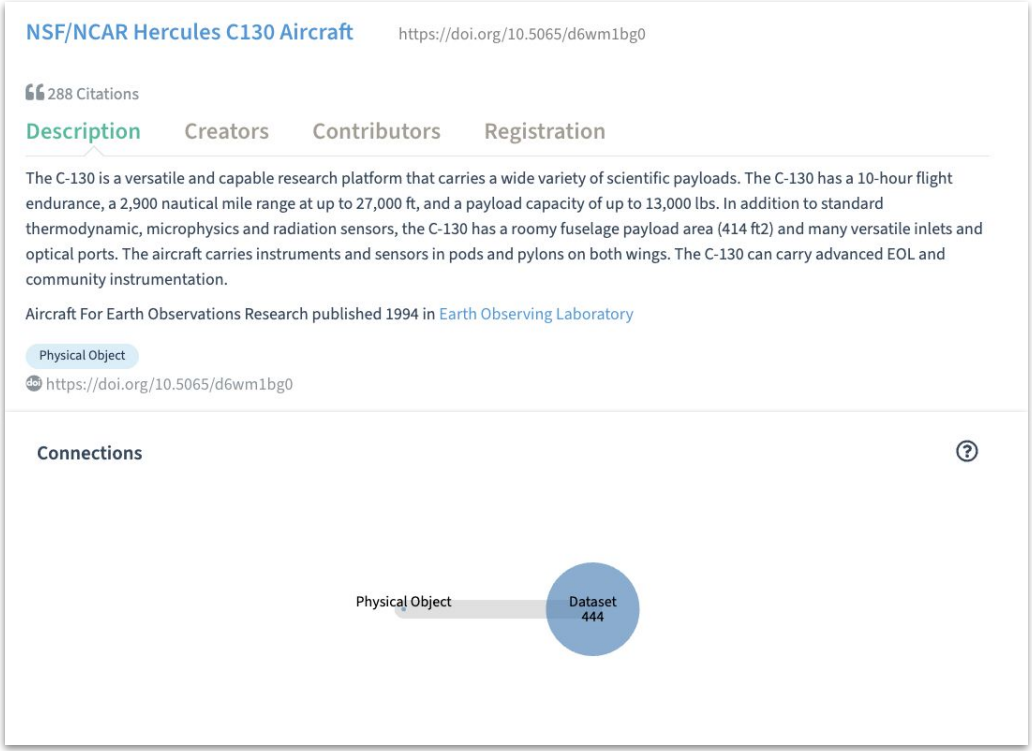
PIDINST Property	DataCite v. 4.5	Comments
Identifier	1. Identifier	
identifierType	1.a identifierType "DOI"	
Name	3. Title	May be the title of a dataset, the name of a piece of software or instrument.
Owner	7. Contributor with 7.a contributorType: HostingInstitution	Can be used for the owner of an instrument, i.e. the institution responsible for the management of the instrument. This may include the legal owner, the operator, or an institute providing access to the instrument. Use the contributorType "HostingInstitution". The instrument owner may also be included in 4. Publisher. [2]
ownerName	7.1 contributorName	
ownerIdentifier	7.4 nameIdentifier	
ownerIdentifierType	7.4.a nameIdentifierScheme	

Tracking and reporting: analysis



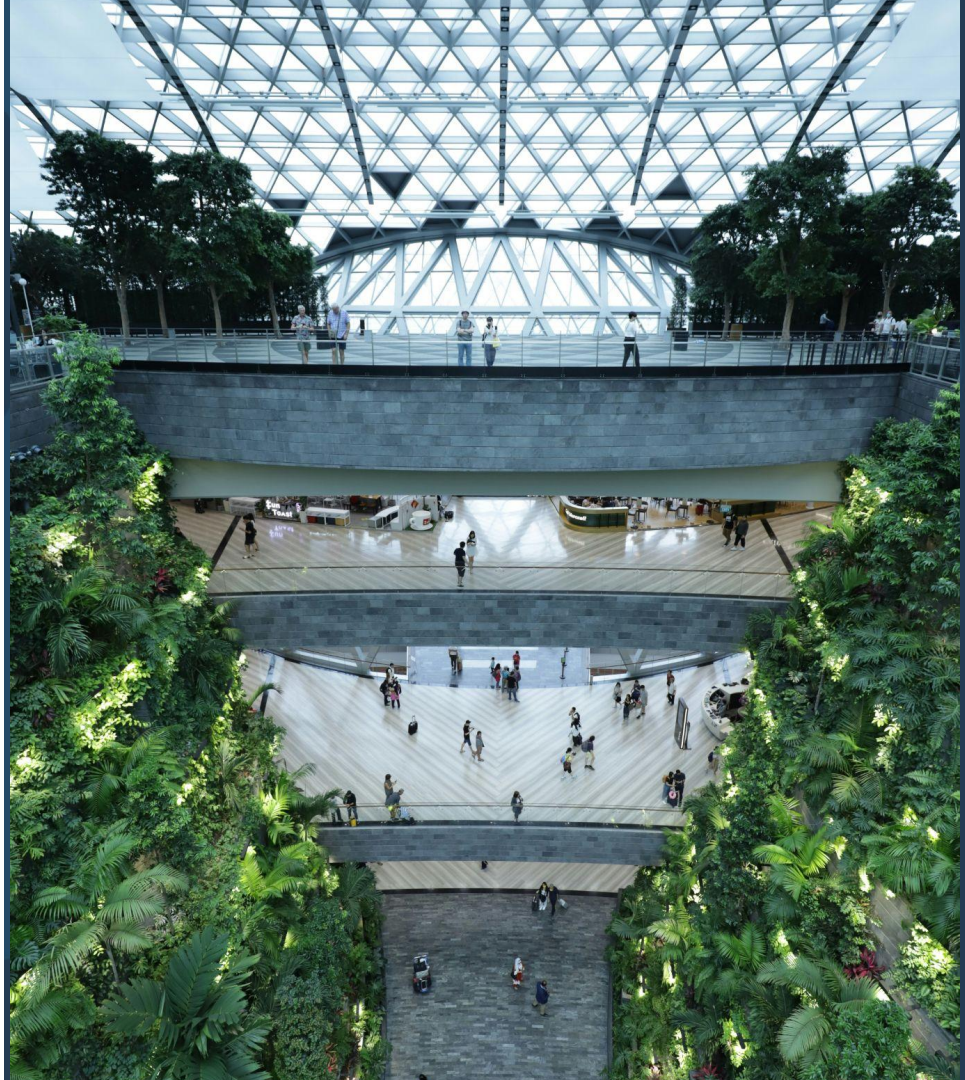
Facility-based tracking and reporting in DataCite Commons discovery interface

Tracking and reporting: citations



Citations reporting in DataCite Commons discovery interface

DataCite alignment with workshop recommendations



Use cases and recommendations: themes and reflections

- **DataCite is already well equipped** to support the use cases and align with recommendations
- **Multiple pathways for alignment** via DataCite services and metadata schema
  - DOI registration
  - Metadata properties
  - Connection metadata
  - Metadata crosswalks/mappings
  - DataCite APIs and discovery interfaces
- **Metadata quality and completeness** is a collective challenge and responsibility!

Use cases
as supported by DataCite

Use case	DataCite support
Traceability	- "Instrument" resource type - Connection metadata to assert relationships and citations - Citations and usage data
Equity	- Metadata openly available via APIs, DataCite Commons, public data file - Metadata harvested by downstream discovery services
Attribution	- Metadata about people and organizations associated with facilities and instruments - Citations and usage data
Reproducibility	- Citation of specific instruments in metadata
Provenance	- Connection metadata to assert relationships and citations - Related identifiers

Granularity recommendations as supported by DataCite

DataCite

Use case	Recommendation	DataCite support
Traceability	Less granular PID assignment (e.g. a small number of PIDs assigned at to high-level entities, such as facilities or instruments as a whole)	- Resource type for “Instrument”
Equity		- Openly available metadata
Attribution		- Metadata properties for creator, affiliation, and funder metadata
Reproducibility	More granular PID assignment (e.g. more PIDs assigned to specific low-level instruments, and potentially to particular instrument configurations, with versioning of PIDs and instruments)	- Related identifiers
Provenance		- Connection metadata to assert relationships and citations - Related identifiers

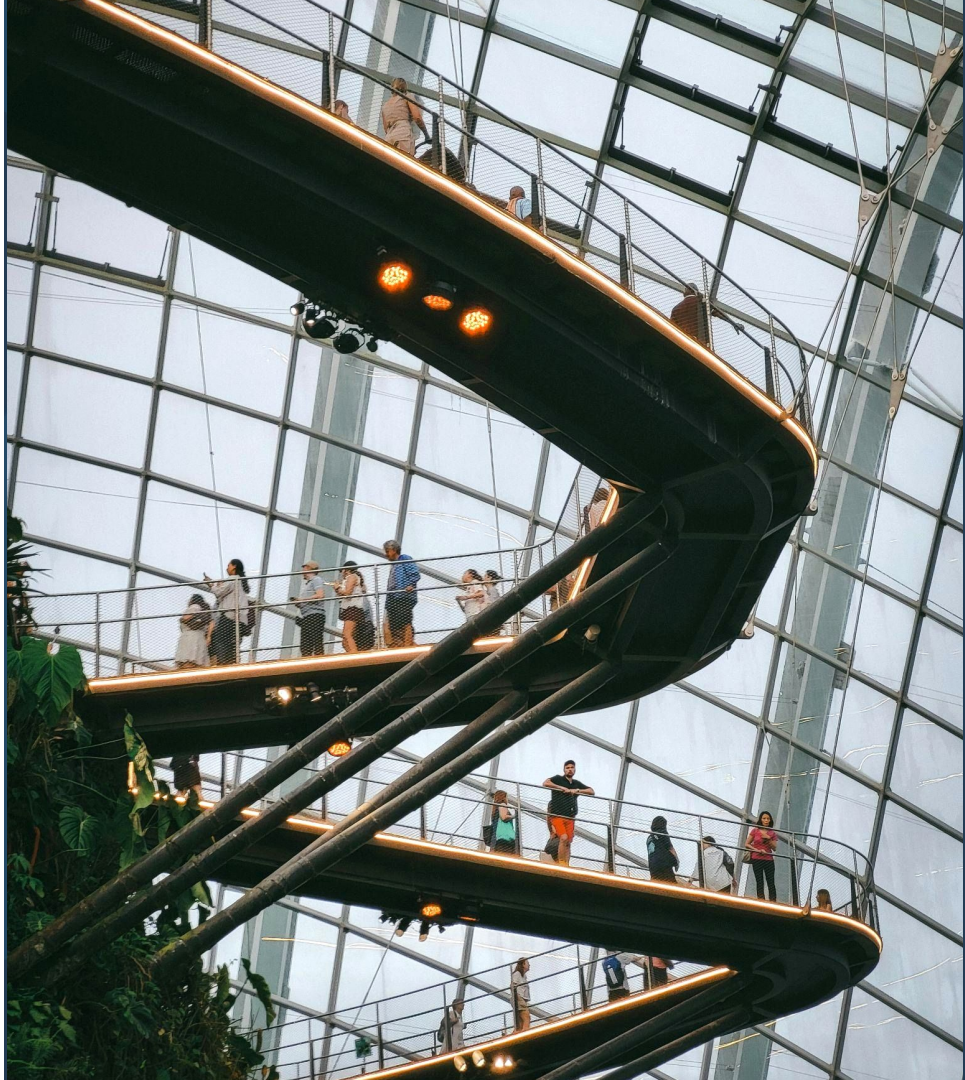
Metadata recommendations as supported by DataCite

Recommendation	DataCite support
PID metadata should align with the PIDINST metadata guidelines	PIDINST schema mapping
High-level metadata for the Traceability, Equity, & Attribution use cases	“Instrument” resource type
For the Reproducibility & Provenance use cases, more detailed platform/instrument metadata is necessary to include in PID and associated metadata (e.g., protocols, dataset metadata)	More detailed metadata (beyond the scope of the DataCite Metadata Schema) can always be linked using a RelatedIdentifier with relationType="HasMetadata"
Associated metadata beyond what can be included in PID metadata, this should be provided by the responsible party and be made available at (or linked from) the PID landing page	More detailed metadata (beyond the scope of the DataCite Metadata Schema) can always be linked using a RelatedIdentifier with relationType="HasMetadata"
Connections between related PIDs should be included in PID metadata whenever they exist.	relatedIdentifiers metadata property supports multiple types of identifiers, including RRIDs

PID provider recommendations as supported by DataCite

Recommendation	DataCite support
Ensure interoperability among facility/instrument PID systems	- PIDINST mapping - Related identifier and connection metadata
Develop services that promote use and tracking of PIDs facilities/instruments	- Citations and metrics
Provide PID metadata options that align with the PIDINST metadata guidelines	- PIDINST mapping
Provide PID metadata options and/or guidance to enable machine-actionable identification of the type of entity to which the PID refers	- Controlled list of resource types - PID metadata in specific properties (ORCID, ROR, related identifiers)

Share your thoughts & get involved



What you can do

- If you're registering DOIs with DataCite: optimize metadata quality and completeness upon registration or update DOIs post-registration
- Share how you're already using DataCite for instrument and facility use cases
- Let us know how DataCite can (better) support your use cases for instruments and facilities?
- Work with us as we explore innovative pathways to metadata enrichments
- Join our community if you're not part of it already!

DataCite

CONNECTING RESEARCH,
ADVANCING KNOWLEDGE

info@datacite.org

datacite.org
datacite.org/blog

support.datacite.org
support@datacite.org

@datacite

@datacite

@datacite

DataCite