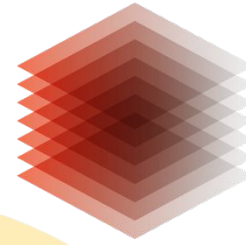


LEIBNIZ-INFORMATIONSZENTRUM
TECHNIK UND NATURWISSENSCHAFTEN
UNIVERSITÄTSBIBLIOTHEK



TIB

PiD

Persistent Identification of Instruments: The RDA WG work so far and a road ahead

Markus Stocker, Rolf Krahl, Louise Darroch, Robert Huber, Ted Habermann, et al.
September 15, 2023
FAIR Facilities and Instruments Workshop, Boulder, Colorado

Motivation

- Instruments play an essential role in creating research data
- Instrument metadata is required to assess data quality and reuse

Borgman (*): “To interpret a digital dataset, much must be known about the hardware used to generate the data, whether sensor networks or laboratory machines.”

- Persistent interlinking with instrument models, platforms, data, facilities, institutions, etc.
- Cite instruments in scientific literature
- Inventory, funding, etc.

Origins

- PIDapalooza, November 2016
- Research Data Alliance, since 2017
 - RDA PIDINST WG endorsement and kick-off, March 2018
 - DSJ Article as RDA Supporting Outputs, July 2020
 - PIDINST metadata schema as RDA Recommendation, March 2022
 - <https://pidinst.org>

Approach

- Persistent identifier for instrument instances used in research
- Use case driven metadata schema development and revisions (community feedback)
- Schema implementation by *existing* PID providers (so far, DataCite, ePIC; possibly, Wikidata)
- Adoption by institutional instrument providers

Schema



Name	Name by which the instrument instance is known
Owner	Institution(s) responsible for the management of the instrument
Manufacturer	The instrument's manufacturer(s) or developer
Model	Name of the model or type of device as attributed by the manufacturer
Description	Technical description of the device and its capabilities
InstrumentType	Classification of the type of the instrument
MeasuredVariable	The variable(s) that this instrument measures or observes
Date	Dates relevant to the instrument
RelatedIdentifier	Identifiers of related resources
AlternatIdentifier	Other identifiers pertaining to the same instrument instance

Schema implementations

- DataCite
 - Based on a PIDINST-DataCite Schema partial mapping
 - Missing attributes, e.g. measured variables, model name
 - Not applicable attributes, e.g. publisher, publication year
 - Bending semantics, e.g. creator for manufacturer
 - With version 4.5 new resource type “Instrument” (release planned for Q4/2023)
 - Global PID provider
- ePIC
 - Full PIDINST schema implementation
 - Less known, European-centric PID provider

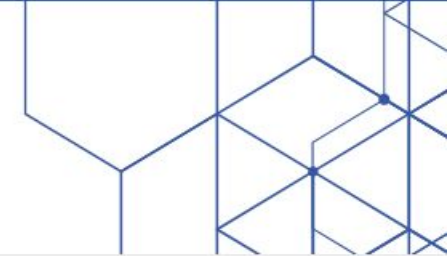
GO TO EUDAT WEBSITE



SEARCH

HELP COMMUNITIES REGISTER CONTACT

Login



Register and publish your scientific instruments

Search for scientific instruments or register as a user to register and publish your own instrument!

[Login or Register](#)

Register instrument

Register a new instrument

<https://b2inst.gwdg.de/>

Metadata Schema for the Persistent Identification of Instruments

RDA Recommendation



DOI: [10.15497/RDA00070](https://doi.org/10.15497/RDA00070)

Authors: Rolf Krahl, Louise Darroch, Robert Huber, Anusuriya Devaraju, Jens Klump, Ted Habermann, Markus Stocker, RDA PIDINST WG members

Published: 30 March 2022

Version: 1.0

Abstract: Instruments play an essential role in creating research data but they are often only identified in scientific literature using free text. Through the use of globally unique, persistent identifiers (PIDs) such as the Digital Object Identifier (DOI), it is now common practice to establish traceable links between scientific literature and the data that generated it. Such cross-linking has received considerable attention in the community in recent years and has been generalized to other entities, including people, organizations, funders, etc. Given the importance of instruments and associated metadata to the assessment of data quality and data reuse, globally unique, persistent and resolvable identification of instruments is also crucial. The Research Data Alliance Persistent Identification of Instruments Working Group (PIDINST WG) developed and demonstrated a solution for publishing persistent identifiers for instruments at established PID providers. This solution comprises a metadata schema for the persistent identification of instruments (PIDINST Schema).

Schema is the recommended schema for the persistent identification of instruments or bodies using or managing instruments. The Schema is PID provider tested with ePIC and DataCite by numerous institutions including Institut für Materialien und Energie GmbH, British Oceanographic Data Centre,

<https://doi.org/10.15497/RDA00070>

License: [Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

RDA webpage: <https://www.rd-alliance.org/group/persistent-identification-instruments-wg/outcomes/metadata-schema-persistent-identification>

Citation and Download: Krahl, R., Darroch, L., Huber, R., Devaraju, A., Klump, J., Habermann, T., Stocker, M., & The Research Data Alliance Persistent Identification of Instruments Working Group members. (2021). Metadata Schema for the Persistent Identification of Instruments. *Research Data Alliance*. <https://doi.org/10.15497/RDA00070>

Outputs



The screenshot shows the PIDINST documentation website. The left sidebar contains a search bar and a list of documents: 'PIDINST White Paper' and 'ePIC Cookbook'. The main content area features the title 'Persistent Identification of Instruments' with a 'PID' logo. Below the title is a paragraph explaining the group's mission. A 'Measuring instruments' section follows, describing their use in various scientific fields. The 'PIDINST is a working group' section mentions its affiliation with the Research Data Alliance (RDA). A 'The group produced the following outputs:' section lists a paper in Data Science Journal and the 'ePIC Cookbook'. The page footer includes a Creative Commons Attribution 4.0 International License logo and a 'Next' button.

Docs » Persistent Identification of Instruments [Edit on GitHub](#)

Persistent Identification of Instruments

The [Persistent Identification of Instruments WG \(PIDINST\)](#) seeks to explore a community-driven solution for globally unique identification of measuring instruments operated in the sciences.

Measuring instruments, such as sensors used in environmental science, DNA sequencers used in life sciences or laboratory engines used for medical domains, are widespread in most fields of applied sciences. The ability to link an active instrument (instance) with an instrument type and with the broader context in which the instrument operates, including generated data, other instruments and platforms, people and manufacturers, etc., is critical, especially for automated processing of such contextual information and for the interpretation of generated data.

PIDINST is a working group in the [Research Data Alliance \(RDA\)](#). It aims to establish a cross-discipline, operational solution for the unique and lasting identification of measuring instruments actively operated in the sciences.

The group produced the following outputs:

- [Persistent Identification of Instruments](#). *Data Science Journal*, 19: 18, pp. doi.org/10.5334/dsj-2020-018 provides an overview of the work of PIDINST.
- [ePIC Cookbook](#). Detailed instructions on how to create instrument PIDs using the ePIC infrastructure.

The content of this site is licensed under a [Creative Commons Attribution 4.0 International License](#).

[Next](#)

<https://docs.pidinst.org/>

Outputs



 **DATA SCIENCE JOURNAL**

Reading: Persistent Identification of Instruments Share: [f](#) [t](#) [g+](#) [in](#)

Special Collection: [Research Data Alliance Results](#)

Research Papers

Persistent Identification of Instruments

Authors: [Markus Stocker](#) , [Louise Darroch](#), [Rolf Krahl](#), [Ted Habermann](#), [Anusuriya Devaraju](#), [Ulrich Schwardmann](#), [Claudio D'Onofrio](#), [Ingemar Häggström](#)

Abstract

Instruments play an essential role in creating research data. Given the importance of instruments and associated metadata to the assessment of data quality and data reuse, globally unique, persistent and resolvable identification of instruments is

Working Group Persistent Identification of a community-driven solution for persistent we present and discuss in this paper. Based on an developed a metadata schema and prototyped Cite and ePIC as representative persistent identifier infrastructures and with HZB (Helmholtz-Zentrum Berlin für Materialien und Energie) and BODC (British Oceanographic Data Centre) as representative institutional instrument providers. These implementations demonstrate the viability of the proposed solution in practice. Moving forward, PIDINST will further catalyse adoption and consolidate the schema by addressing new stakeholder requirements.

Keywords: [Persistent Identification](#), [Instruments](#), [Metadata](#), [DOI](#), [Handle](#)

<https://doi.org/10.5334/dsj-2020-018>

Adoption



Adoption?

- Substantial interest by institutional instrument providers but little concrete adoption, so far
- To our knowledge, few instruments are identified using the proposed approaches
- Why? Possibly because
 - Lacking identifier governance
 - RDA PIDINST WG is not a governing body (and lacks the resources)
 - Some technology challenges with PID providers
 - Difficult to build trust in the identifier

PIDINST-Future

- PIDINST-Future as a response to this problem
- German-led consortium developed the project proposal
- Now submitted to the German Research Foundation (DFG)
- Partners
 - TIB - Leibniz Information Centre for Science and Technology
 - University of Bremen (PANGAEA)
 - Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)
 - Helmholtz Centre Berlin for Materials and Energy (HZB)

PIDINST-Future

- PIDINST COM: Establish a permanent PIDINST Forum
 - Stakeholder engagement, consultation, training and dissemination
 - Inform PIDINST identifier governance
 - Support the testing and validation of solutions
- PIDINST GOV: Establish a governance for the identifier
 - Suitable non-profit organizational form
 - Governing organization will be responsible of future development
 - Membership to ensure organizational sustainability
- PIDINST TEC: Further advance the technical infrastructure
 - Federated search and resolution service (over heterogeneous PID providers)
 - Establish B2INST as low-barrier service
 - Crosslinking instruments

Takeaways

- Foundations for persistent identification of instruments are in place
- Approaches have been tested and validated
- Time to scale adoption
- Sustained international governance for the identifier, build trust
- Hopefully with seed funding with a DFG funded project