COFFEE LECTURE

Fantastic PIDs and how they help you to find things

Nina Weisweiler & Antonia C. Schrader
Helmholtz Association – Helmholtz Open Science Office
WWW, February 02, 2021
Introduction

What are PIDs (Persistent Identifiers)?

▪ “unique universal persistent identifiers” (Koster, 2020)
  ○ “Identifier” = a string of digits referring to an object
  ○ “Unique” = only refers to one object (within the known universe)
  ○ “Universal” = is valid for the whole of the world (or world wide web)
  ○ “Persistent” = shall remain available independent of individual institutions, systems or system implementations

▪ A system of digital identifiers that can refer to digital, physical or abstract objects.
  → However, they are typically used for digital, online accessible objects.

▪ Mostly “actionable”: You can convert them into a URL that resolves to the identified source or to a landing page with metadata information.
How do PIDs work?

- PIDs are usually created within institutionally administered systems.
  - Persistence is purely a matter of service!
    → It always takes someone who commits to resolving them for users.

- PID magic doesn’t emerge from the identifier strings themselves… But from open scholarly infrastructures that provide accompanying metadata and associated services like APIs for getting the metadata.

- The PID string is built according to a consistent schema.
  - Example: DOI have the format prefix/suffix (prefix: identifies the registrant of the identifier / suffix: chosen by the registrant and identifies the specific object associated with that DOI)
  - A DOI formatted as URL: https://doi.org/10.2312/os.helmholtz.017

<table>
<thead>
<tr>
<th>Domain/PID-URI</th>
<th>Prefix</th>
<th>Suffix</th>
</tr>
</thead>
</table>

Introduction
Why are they fantastic?

- PIDs **prevent “link rot”** (= URL does not resolve anymore due to a resource being relocated or becoming permanently unavailable).
- PIDs provide **unique identification** for publications, datasets, persons, organizations, research material, and much more...
- Especially in academia, they provide useful standards for **unambiguous citation**.
- PIDs can be used to unambiguously **link to other PIDs**.
  - E.g. linking articles in journal citations or researcher and their datasets
- PIDs facilitate the comprehensive and correct **assignment of research contributions** (publications, datasets, conference contributions etc.) to the scientific record.
- The use of PIDs is **recommended** by important research funders and organisations.
Overview PID-Landscape

<table>
<thead>
<tr>
<th>Research entity</th>
<th>PID types used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication</td>
<td>DOI, Accession number, <a href="#">Handle</a>, <a href="#">URN</a>, Web of Science UID, PMID, PMC, arXiv Identifier, BibCode, ISSN, ISBN, <a href="#">PURL</a></td>
</tr>
<tr>
<td>Data</td>
<td>DOI, Accession number, Handle, PURL, URN, <a href="#">ARK</a></td>
</tr>
<tr>
<td>Researcher (or Scholar)</td>
<td><a href="#">ORCID</a> iDs, GND, ISNI, ResearcherIDs, ScopusIDs</td>
</tr>
<tr>
<td>Citation</td>
<td><a href="#">Open Citation Identifier (OCI)</a></td>
</tr>
<tr>
<td>Conference</td>
<td>DOI, Accession number</td>
</tr>
<tr>
<td>Organization</td>
<td>GRID, Ringgold iDs, ROR iDs</td>
</tr>
</tbody>
</table>

→ [Overview of services related to Persistent Identifiers (PIDs)](#)
## Overview PID-Landscape

<table>
<thead>
<tr>
<th>Research entity</th>
<th>PID types used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrument, Device, Sensor, Platform, Research Facility</td>
<td>DOI, <a href="https://example.com">RRID</a>, UID</td>
</tr>
<tr>
<td>Archival/Storage facility</td>
<td>URI, DOI, UUID</td>
</tr>
<tr>
<td>Geological or Biological Sample</td>
<td>Accession number, RRID, DOI, IGSN</td>
</tr>
<tr>
<td>Project</td>
<td>local identifier, accession number, <a href="https://example.com">RAiD</a></td>
</tr>
<tr>
<td>Experiment</td>
<td>/</td>
</tr>
<tr>
<td>Data repository</td>
<td>/ (see re3data COREF)</td>
</tr>
</tbody>
</table>

→ [Overview of services related to Persistent Identifiers (PIDs)](https://example.com)
ORCID

- ORCID stands for **Open Researcher and Contributor ID**

- The ORCID iD is an unique, persistent identifier free of charge to researchers
  - Example: [http://orcid.org/0000-0003-3334-2771](http://orcid.org/0000-0003-3334-2771)

- Your ORCID iD distinguishes you from every other researcher in the world.
Why ORCID?

Solving the name ambiguity issue

One of the crucial reasons we need @ORCID_Org

<table>
<thead>
<tr>
<th>ORCID ID</th>
<th>First name</th>
<th>Family name</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="https://orcid.org/0000-0002-3185-4258">https://orcid.org/0000-0002-3185-4258</a></td>
<td>Thomas</td>
<td>Müller</td>
</tr>
<tr>
<td><a href="https://orcid.org/0000-0002-6914-9570">https://orcid.org/0000-0002-6914-9570</a></td>
<td>Thomas</td>
<td>Müller</td>
</tr>
<tr>
<td><a href="https://orcid.org/0000-0002-6971-5747">https://orcid.org/0000-0002-6971-5747</a></td>
<td>Thomas</td>
<td>Müller</td>
</tr>
<tr>
<td><a href="https://orcid.org/0000-0001-8726-6052">https://orcid.org/0000-0001-8726-6052</a></td>
<td>Thomas</td>
<td>Müller</td>
</tr>
<tr>
<td><a href="https://orcid.org/0000-0002-0013-2156">https://orcid.org/0000-0002-0013-2156</a></td>
<td>Thomas</td>
<td>Müller</td>
</tr>
<tr>
<td><a href="https://orcid.org/0000-0002-2857-428X">https://orcid.org/0000-0002-2857-428X</a></td>
<td>Thomas</td>
<td>Müller</td>
</tr>
<tr>
<td><a href="https://orcid.org/0000-0002-3400-3248">https://orcid.org/0000-0002-3400-3248</a></td>
<td>Thomas</td>
<td>Müller</td>
</tr>
<tr>
<td><a href="https://orcid.org/0000-0003-1427-2982">https://orcid.org/0000-0003-1427-2982</a></td>
<td>Thomas</td>
<td>Müller</td>
</tr>
<tr>
<td><a href="https://orcid.org/0000-0001-5331-5522">https://orcid.org/0000-0001-5331-5522</a></td>
<td>Thomas</td>
<td>Müller</td>
</tr>
<tr>
<td><a href="https://orcid.org/0000-0002-7646-8115">https://orcid.org/0000-0002-7646-8115</a></td>
<td>Thomas</td>
<td>Müller</td>
</tr>
</tbody>
</table>
ORCID

- Other author/person identifier:
  - International Name Standard Identifier (ISNI)
  - Gemeinsame Normdatei (GND) – in English: the integrated authority file
  - ResearcherID (Publons)
  - Scopus Author ID

- Over 10m researchers worldwide have an ORCID iD
  - Over 220k researchers in Germany have an ORCID iD

- Provider: international consortium - not-for-profit organization
  - orcid.org
  - German ORCID consortium

- The DFG-funded project ORCID DE supports the distribution of ORCID iDs and the implementation of ORCID in Germany.
ROR

- ROR stands for Research Organization Registry
- ROR is a community-led project to develop an open, sustainable, usable, and unique identifier for every research organization in the world → ROR ID
- ROR focused on being a top-level registry of organizations

- Other OrgIDs:
  - ISNI
  - The integrated authority file (GND)
  - GERiT (ID by the Deutsche Forschungsgemeinschaft (DFG))
  - GRID
  - Ringgold ID
Namensvariationen #1

- Helmholtz-Zentrum Potsdam - Deutsches GeoForschungsZentrum GFZ
- GeoForschungsZentrum (Deutsches GeoForschungsZentrum)
- Helmholtz-Zentrum (Potsdam)
- Helmholtz Centre (Potsdam)
- German Research Centre for Geosciences
- Research Centre for Geosciences (Potsdam)
- Hermann-von-Helmholtz-Gemeinschaft Deutscher Forschungszentren. Helmholtz-Zentrum (Potsdam)
- Hermann-von-Helmholtz-Gemeinschaft Deutscher Forschungszentren. Helmholtz Centre (Potsdam)
- Hermann-von-Helmholtz-Gemeinschaft Deutscher Forschungszentren. Deutsches GeoForschungsZentrum
- Hermann-von-Helmholtz-Gemeinschaft Deutscher Forschungszentren. German Research Centre for Geosciences
- GFZ (Abkürzung)
- Deutsches GeoForschungsZentrum GFZ
ROR

- ROR stands for Research Organization Registry.
- ROR is a community-led project to develop an open, sustainable, usable, and unique identifier for every research organization in the world.
- The ROR ID is an organization identifier (OrgID) free of charge that focuses on being a top-level registry of organizations.

- Other OrgIDs:
  - ISNI
  - The integrated authority file (GND)
  - GERiT (ID by the Deutsche Forschungsgemeinschaft)
  - GRID
  - Ringgold ID

- The ORCID DE project provides regular information on the developments of ROR on its blog.
  - More information: https://ror.org/
Some findings from the re3data COREF PIDapalooza session:

- **What could be the use cases for repository PIDs?**
  - A. Repository PID might be useful in dataset citations, esp. for monitoring of dataset usage and dissemination.
  - B. Repository PID might be useful to avoid being dependent on verified repository names.
  - C. Repository PID might allow the identification and mapping of repositories across different repository registries and would therefore aid researchers and infrastructure providers. (e.g. DataCite Commons)
  - D. “In creating PID Graphs of research outputs, we want to be able to store the relationships between datasets and repositories as triplets. This requires repository PIDs.”

- **Further recommendations:**
  - Get one system for both research data repositories and text-based publication repositories.
  - A repository PID should refer to a landing page with metadata and not directly to the repository website.
KEEP IN TOUCH AND STAY INFORMED

- Mail – open-science@helmholtz.de
- Website – www.os.helmholtz.de
- Make sure to sign up for our Open Science Newsletter
- Mailing list – Helmholtz Open Science Professionals
- Twitter – @helmholtz_os
- Make use of the Helpdesk provided by OA-network
KEEP IN TOUCH AND STAY INFORMED

- Mail – open-science@helmholtz.de
- Website – www.os.helmholtz.de
- Make sure to sign up for our Open Science Newsletter
- Mailing list – Helmholtz Open Science Professionals
- Twitter – @helmholtz_os
- Make use of the Helpdesk provided by OA-network

57th Helmholtz Open Science Online Seminar

Practical Steps Towards Open and Reproducible Research

Wednesday, February 10, 2021, with Dr. Heidi Seibold

Registration
QUESTIONS & ANSWERS
MANY THANKS FOR YOUR PARTICIPATION!

Nina Weisweiler  
https://orcid.org/0000-0001-6967-9443

Antonia C. Schrader  
https://orcid.org/0000-0001-7080-634X

All texts in this presentation, except citations, are licenced under Attribution 4.0 International (CC BY 4.0):  
https://creativecommons.org/licenses/by/4.0/deed.de
LITERATURE

- YouTube (2018): IGSN: Assign a persistent identifier to physical samples. https://www.youtube.com/watch?v=Et7q6EOGNc&ab_channel=AustralianResearchDataCommons-ARDC
- OpenAIRE: https://www.openaire.eu/what-is-a-persistent-identifier
- Australien Reserach Data Commons: https://ardc.edu.au/services/identifier/
- ...