COFFEE LECTURE

**Fantastic PIDs and
how they help you to find things**

Nina Weisweiler & Antonia C. Schrader
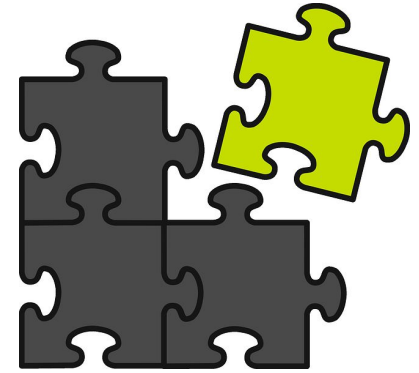
Helmholtz Association – Helmholtz Open Science Office

WWW, February 02, 2021

Introduction
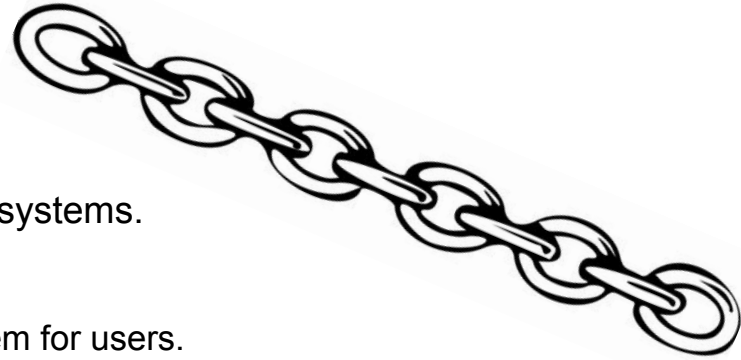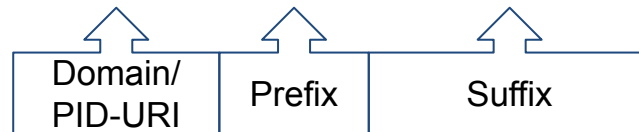
What are PIDs (Persistent Identifiers)?

- "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.

Introduction

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

Domain/ PID-URI	Prefix	Suffix

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

Explore the _PID Graph_

- 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

Research entity	PID types used
Publication	DOI, Accession number, Handle, URN, Web of Science UID, PMID, PMC, arXiv Identifier, BibCode, ISSN, ISBN, PURL
Data	DOI, Accession number, Handle, PURL, URN, ARK
Researcher (or Scholar)	ORCID iDs, GND, ISNI, ResearcherIDs, ScopusIDs
Citation	Open Citation Identifier (OCI)
Conference	DOI, Accession number
Organization	GRID, Ringgold iDs, ROR iDs

→ Overview of services related to Persistent Identifiers (PIDs)

Overview PID-Landscape

Research entity	PID types used
Instrument, Device, Sensor, Platform, Research Facility	DOI, RRID, UID
Archival/Storage facility	URI, DOI, UUID
Geological or Biological Sample	Accession number, RRID, DOI, IGSN
Project	local identifier, accession number, RAiD
Experiment	/
Data repository	/ (see re3data COREF)

→ Overview of services related to Persistent Identifiers (PIDs)

ORCID

- ORCID stands for **O**pen **R**esearcher and **C**ontributor **ID**

- The ORCID iD is an unique, persistent identifier free of charge to researchers
  - Example: 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

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

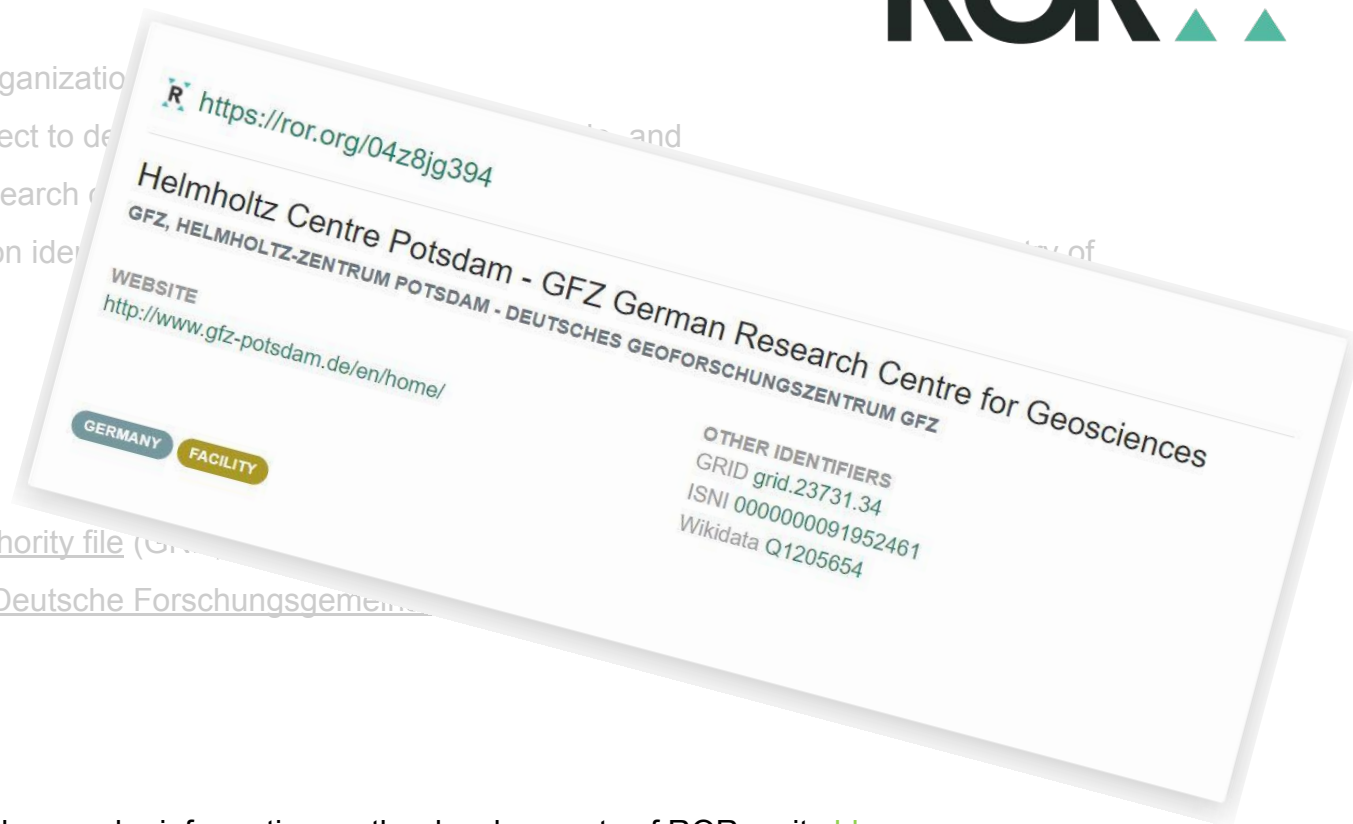- ROR stands for **R**esearch **O**rganization **R**egistry
- ROR is a community-led project to develop an open, sustainable, usable, and unique identifier for every research organization in the world → ROR ID
- ROR focused on being a top-level registry of organizations

- 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 **R**esearch **O**rganizatio...
- ROR is a community-led project to de... and
  unique identifier for every research...
- The ROR ID is an organization ide... ...ty of
  organizations

- Other OrgIDs:
  - ISNI
  - The integrated authority file (G...
  - GERiT (ID by the Deutsche Forschungsgemein...
  - GRID
  - Ringgold ID

- The ORCID DE project provides regular information on the developments of ROR on its blog.
  - More information: https://ror.org/

re3data COREF – PIDs for research (data) repositories

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.

HELMHOLTZ
Open Science

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

HELMHOLTZ
Open Science

KEEP IN TOUCH AND STAY INFO...

- Mail – open-science@helmholtz.de

- Website – www.os.helmholtz.de

- Make sure to sign up for our Open Science New...

- Mailing list – Helmholtz Open Science Professional...

- 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

HELMHOLTZ
Open Science

MANY THANKS FOR YOUR PARTICIPATION!

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

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

LITERATURE

- Koster (2020): Persistent identifiers for heritage objects. In: Code4Lib Journal, 47, https://journal.code4lib.org/articles/14978

- Ferguson, Christine, et al. (2018): D3.1 Survey of Current PID Services Landscape (Version 1). In: Zenodo.

  https://doi.org/10.5281/zenodo.1324296

- Vierkant, Bertelmann (2020): Helmholtz Open Science Office - Shaping the PID Landscape in Germany and Beyond - Contributions,

  PIDapalooza (Lisbon, Portugal 2020). https://doi.org/10.6084/m9.figshare.11822487

- Plomp (2020): Going digital. Persistent Identifiers for research samples, resources and instruments. *Data Science Journal, 19*(1), 46.

  https://doi.org/10.5334/dsj-2020-046

- Dreyer, Hagemann-Wilholt, et al. (2019): Die Rolle der ORCID iD in der Wissenschaftskommunikation: Der Beitrag des

  ORCID-Deutschland-Konsortiums und das ORCID-DE-Projekt. In: ABI Technik. https://doi.org/10.1515/abitech-2019-2004

- Kraft, Dreyer (2018): Do researchers need to care about PID systems?. Presented at the PIDapalooza 2018, Girona, Spain: Zenodo.

- Cousijn, Braukmann, et al. (2021). Connected research: the potential of the PID graph. *Patterns*, *2*(1), 100180.

  https://doi.org/10.1016/j.patter.2020.100180

- YouTube (2018): IGSN: Assign a persistent identifier to physical samples.

  https://www.youtube.com/watch?v=Ei7gp6EOGNc&ab_channel=AustralianResearchDataCommons-ARDC

- OpenAIRE: https://www.openaire.eu/what-is-a-persistent-identifier

- Australien Reserach Data Commons: https://ardc.edu.au/services/identifier/

- …