Enabling Open Science Through Persistent Identifiers for Research Instruments and Facilities

FAIR Facilities and Instruments
NSF FAIR Open Science (FAIROS)
Research Coordination Network (RCN)

https://ncar.github.io/FAIR-Facilities-Instruments/

Organizations and Project Personnel

NCAR | NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Matt Mayernik
Greg Stossmeister

University of Colorado Boulder

Andrew Johnson
Aditya Ranganath
Matthew Murray

NSF Awards
2226396
2226397
2226398

FLORIDA STATE UNIVERSITY

Renaine Julian

Stanford University

Claudius Mundoma

Project Goals

- Create and facilitate Research Coordination Network focused on the assignment of Persistent Identifiers (PIDs) to research facilities and instrumentation
- Aggregate use cases for why/how PIDs can be assigned to facilities and instruments
- Facilitate generation of expertise and guidance on the key topics of interest
- Provide recommendations and lessons learned targeted to the multiple communities involved

Definitions

- (Digital) Persistent Identifier (PID): "A digital identifier that is globally unique, persistent, machine resolvable and processable, and has an associated metadata schema" (National Science and Technology Council)
  - Research Resource Identifiers: "ID Numbers assigned to help researchers cite key resources (antibodies, model organisms, software projects etc.) in the biomedical literature to improve transparency of research methods"
  - Datacite DOIs: "Digital Object Identifiers provide persistent links that ensure the accessibility of your content even if URLs change, enhancing discoverability and credibility"
- Instrument: "A device used for making measurements, alone or in conjunction with one or more supplementary devices" (International Vocabulary of Metrology). Instruments may have configurations or settings that change over time. Can range in size from small temperature sensors to more complex microscopes or RADAR and LIDAR Devices.
- Facilities: Broader than "instrument", in that facilities also refer to people, and may be tied to a specific physical location. For example, a biotechnology laboratory.

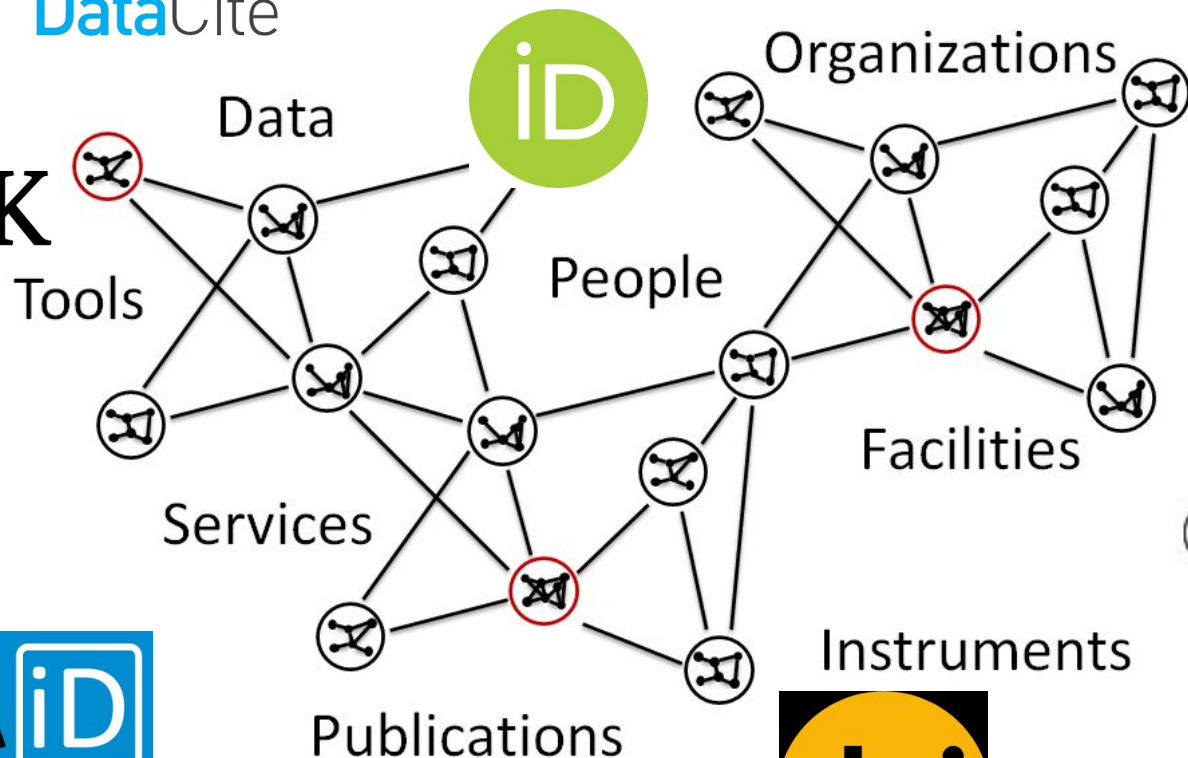# Open Science and PIDs for Facilities and Instruments

- Encouraging the use of PIDs for facilities and instruments advances open science by:
  - Increasing transparency
  - Promoting research reproducibility
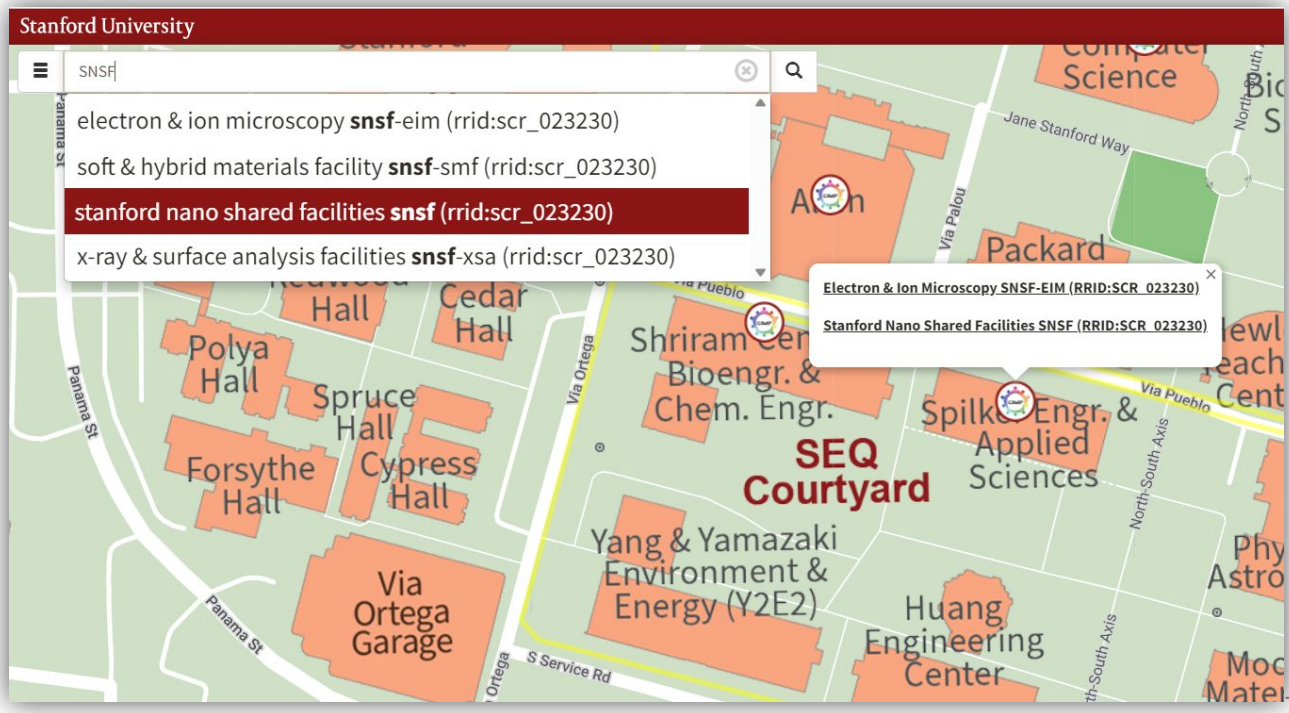 - Fostering equity
  - Supporting discoverability

Networked Science
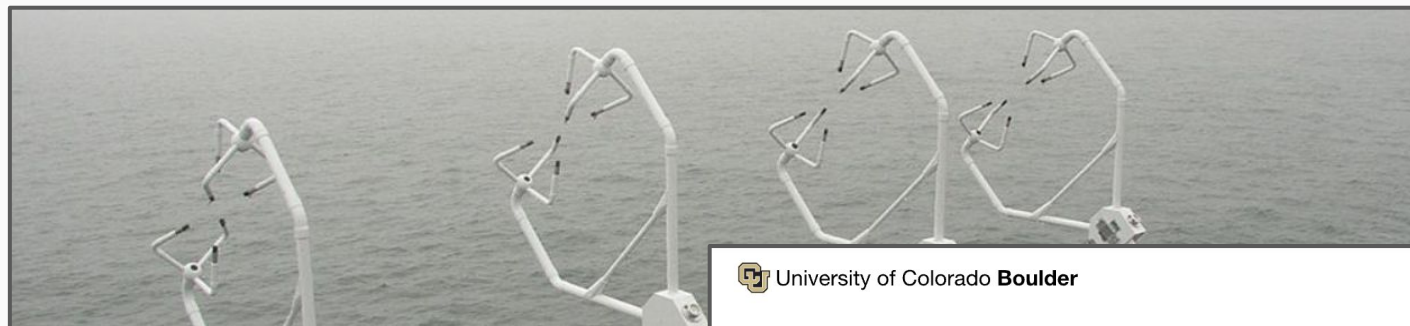
Examples of PID adoption for facilities and instruments

Mapping of Stanford Shared Instrumentation Facilities with RRIDs helps you search, locate, access and cite research capabilities at your fingertips

search,
locate,
access, …
cite

Examples of PID adoption for facilities and instruments

https://doi.org/10.5065/D6ZC80XJ

https://doi.org/10.25811/k3w6-pk81

Year 1 Project Efforts

- Developing [project web site](#)
- Identifying relevant people, projects, information, and resources
- Conference engagement - AMS, ABRF, ESIP
- Focus groups & presentations to relevant groups
  - NSF FAIR Open Science RCN project cohort
  - Focus group discussion with Earth science facility users/providers
  - FSU campus facilities staff
  - CI Compass - FAIR Data Working Group
- In-person workshop in Boulder, CO

Workshop #1 Boulder, CO

- 35 participants from:
  - 17 U.S. states
  - Academic institutions, national labs, nonprofit orgs, publishers, industry
  - Biomedical science, geological science, environmental science, space science, materials science, and more
- [18 presentations](#)
- 3 breakout sessions
- [Workshop report](#)

Observation #1 - Need

There is a need for PIDs for research instrumentation to encourage scientific reproducibility, ensure provenance of data, and provide credit for instrument developers and providers.

Community Responsibilities - All stakeholders should encourage the assignment of PIDs to research instrumentation, as well as the citation of those instruments in scholarly work that results from their use.

Observation #2 - PID Systems

The current use of PIDs for instruments is scattered and inconsistent in how and which PIDs are used. Multiple PID systems are already being used for the purposes of identifying research instrumentation.

Next Step For Our Project - A thorough comparison of the existing PID systems should be conducted, to evaluate their applicability for identification of facilities, platforms, and instruments for the most common use cases.

Observation #3 - Adoption

In general, the question of which PID system to use is less important than finding ways to lower the barrier for adoption of instrument PIDs and better communicate the value of using them.

Next Step For Our Project - Develop guidance on how to navigate adoption considerations for the different PID options.

Community Responsibilities - PID systems should continue to work to make connections, to enable interoperability in services, such as PID resolving and metadata, and to clarify the value of each respective system.

Observation #4 - Metadata

Sometimes we need to think about metadata, in addition to PIDs. PIDs may not be appropriate to solve all challenges related to research traceability, transparency, and reproducibility.

Next Step For Our Project - Working from the [PIDINST](#) metadata recommendations, develop a set of implementation guidelines for the scientific community that includes how to keep metadata relevant and up to date and connect instrument PIDs to resulting dataset metadata.

Observation #5 - Granularity and Evolution

Granularity and evolution considerations can be very complicated for research instrumentation. A good rule of thumb is to start simple, and then move to more complicated approaches if needed.

Next Step For Our Project - Granularity and instrument evolution challenges need to be better understood across the various disciplines involved in order to develop guidelines that will support the highest community priorities whether those are impact tracking, credit to developers/operators and/or scientific reproducibility.

15

Observation #6 - Resource Limitations

Instrument and facility providers often face significant resource limitations that make assigning, managing, and promoting PIDs challenging.

Next Step For Our Project - Continue to gather data from research facilities and assess challenges across disciplines to develop guidance for how instrument and facility providers can most easily and efficiently begin the process of assigning PIDs.

Observation #7 - Value

Convincing users of the value of citing PIDs for instruments will be important to advance adoption.

Next Step For Our Project - Develop value statements about how PIDs help instrumentation and facilities providers to contribute to research being "born FAIR."

Observation #8 - Incentives

While incentives can generally increase PID use and adoption, a lack of understanding of certain stakeholders' incentives can be a challenge. Different incentives are needed on both the researcher and administrative sides to ensure adoption of PIDs.

Next Step For Our Project - Learn more about how PID use and citation can be made easier in hopes of encouraging use. Building better understanding of researcher and PI benefits and incentives for PID use and citation may reduce challenges.
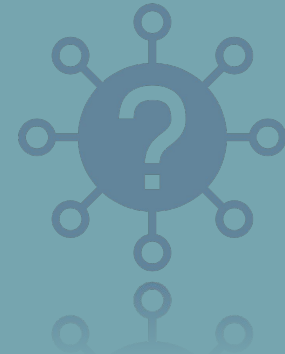
We would love to hear from you!

Project website:

https://ncar.github.io/FAIR-Facilities-Instruments/

Sign-up for project communications here ➡️

Next workshop at Florida State University, August 20-22, 2024 - Details coming soon

Discussion

- What is the relevance of Persistent Identifiers (PIDs) for your professional responsibilities or interests?
- Do you currently assign PIDs to scholarly tools or outputs of any kind? If so, which PID system(s) do you use? Why did you/your institution choose them? How do you decide which scholarly objects receive PIDs?
- What are your main use cases for assigning PIDs to scholarly objects (either currently, or desired for the future)? In other words, what scholarly objects are currently assigned PIDs, and what scholarly objects might be assigned PIDs in the future?
- What challenges are you encountering (or do you foresee) in assigning PIDs to scholarly objects, including research facilities and instruments?
- What might be relevant challenges or constraints that could slow down PID adoption, and are there possible solutions for these challenges or constraints?
- What are community organization or communication strategies that could be used to facilitate the adoption of PIDs to scholarly objects in general, and to research instruments and facilities in particular?