

# Reproducibility through Identifiers: Persistent Identifiers (PIDs) for research facilities and instruments

Matthew Murray, Andrew Johnson, Adi Ranganath, Renaine Julian,  
Matt Mayernik, and Claudius Mundoma



# Collaborative Research: Cross-Cutting Improvements: FAIR Facilities and Instruments: Enabling transparency, reproducibility, and equity through persistent identifiers

- Funded by the American National Science Foundation's FAIR Open Science Research Coordination Network
- NSF Awards #2226396, 2226397, 2226398
- <https://ncar.github.io/FAIR-Facilities-Instruments/>



# Definitions

- Instruments
- Facilities
- PIDS
- FAIR

# What is an 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. They can range in size from small temperature sensors to more complex microscopes or RADAR and LIDAR devices.
- A specific remote sensing unit might be considered an instrument.

# What is a facility?

- It's broader than an "instrument", in that facilities may also include people and may be tied to a specific physical location. For example, a biotechnology laboratory (and its staff!).
- A facility may contain multiple instruments
  - Satellites often contain multiple instruments and generate multiple discrete data streams and might be considered a "facility"

# What are PIDs?

- (Digital) Persistent Identifiers
  - A string of letters and numbers
    - $\text{P}\text{T}\text{T}\text{A}\text{E}^{-4}\text{Ø}_9\text{Ç}$
  - Globally unique
  - Persistent
  - Machine resolvable and processable
  - Have an associated metadata schema

# Examples of PIDs

- RRID
- URN
- PMID
- DOI
- DMPID
- Handles
- ARK
- PURL
- Research Organization Registry IDs
- TST
- WFFL

**Wow, I've never heard of  
some of those.**

- That's because I made some of them up.

# What is FAIR?

- **Findable** - How do we enable people to find relevant facilities or instruments?
- **Accessible**- How do we enable facilities and instruments to be accessible by wider audiences?
- **Interoperable** - How do we consistently capture relationships between persistent identifiers?
- **Reusable** - How can we incorporate information about facilities and instruments into data set provenance metadata more consistently?

**Examples**

Alpine Supercomputing  
Cluster, University of  
Colorado Boulder Research  
Computing  
[https://doi.org/10.25811/k3w  
6-pk81](https://doi.org/10.25811/k3w6-pk81)

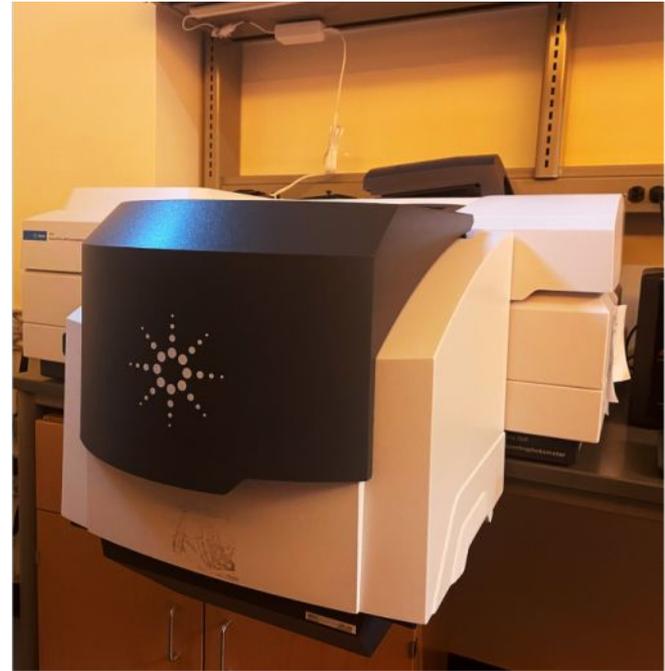


National Center for Atmospheric Research's Lockheed Hercules C130 Aircraft <https://doi.org/10.5065/D6WM1BG0>



# Polymeric and Optical Materials Characterization Shared Facility

RRID:SCR\_022288



- Revolution InstaGLO® R180 Toaster
- (This is not a facility or instrument.)
- (As far as I know.)
- (This is what a TST identifier is for.)



# The Project

# What are our goals?

- Create and facilitate a Research Coordination Network focused on the assignment of Persistent Identifiers to research facilities and instruments
- Gather use cases for why & how PIDs can be assigned to facilities and instruments
- Facilitate conversations for relevant stakeholders on the key topics of interest
- Provide community-based recommendations and best practices on how to use PIDs in different disciplines

# Why are we doing this

- There is, as of yet, no standard for using identifiers for scientific instruments and facilities
- Encouraging the use of PIDs for facilities and instruments advances open science and allows for:
  - Increasing transparency
  - Promoting research reproducibility
  - Fostering equity
  - Supporting discoverability
  - Ensuring provenance of data
  - Providing credit to developers and providers

# What we've done

- Organized focus groups & presentations for relevant stakeholders
- Held an in-person workshop in Boulder, September 2023
  - Attendees from academic institutions, national labs, nonprofit orgs, publishers, and industry
  - Biomedical science, geological science, environmental science, space science, materials science, and more
  - Slide decks from presentations are available [on our website](#)
- Published: [Fair Facilities And Instruments Workshop #1 Report: Exploring Persistent Identifier Needs, Barriers and Incentives](#)
- Presented at IASSIST/CARTO (and other events!)

**What we've learned**

# What we've learned

- There is a need for these to exist
  - People want them! They think they'd be useful.
- Nobody knows which ones to use
  - There are many different PID systems and there is no “one size fits all” PID for all instruments and facilities
  - Which PID system to use is less important than finding ways to lower the barrier for adoption of instrument PIDs
- PIDs won't solve every problem

# Barriers

# Why don't/won't people use them?

- They don't have identifiers
  - Nobody wants to register their toaster
- They don't know they exist
  - There are many labs and instruments that have PIDs but for whatever reasons researchers don't know about them
- They don't think they're valuable
  - We need to convince researchers of the value
- The PID is printed on the bottom of the toaster and nobody wants to flip it over to get it because they'll get crumbs everywhere.

# Incentives

# Why should people use them?

- **Reproducibility and Replicability**
  - Know what settings were used to make the toast you read about in a journal article
- **Data provenance**
  - Where did this toast come from?
- **Attribution: Track impact and citations**
  - Know who's using the toaster your lab spent \$10,000 on
- **Discoverability and Collaboration: Find and share resources**
  - Not everyone needs to have their own toaster in their office

# Questions

# Who's going to issue all of these PIDs?

- Instrument and facility providers often face significant resource limitations that make assigning, managing, and promoting PIDs challenging.
- How can we ensure PIDs are created and up to date?

# When and how should they be used?

- What metadata needs to be included?
  - Colour of the toaster vs. how many slots
- At what granularity should PIDs be assigned?
  - Does every element/configuration of an instrument need it's own PID?
  - Do you need a general PID for the toaster or do you need a PID specifically for bagel mode?
- The Scientific Instrument of Theseus
  - Instruments and facilities evolve over time
  - When is a new PID issued vs. metadata updated?
  - New software? New hardware?

# Suggestions

# Changes to facilitate PID adoption

- **Culture change:** It is important to create a widespread norm around the citation of research instruments. Creating a cultural norm around instrument/facility citation in relevant scholarly communities will take time, but we can begin to catalyze it through small steps.
  - GUILT: Guiding Users Into Linking Texts
- **Policy change:** Change will also require relevant stakeholders (such as funders, journals, and universities) to adopt policies that encourage the citation of instruments and facilities.
- **Institutional change:** Relevant institutions must change and further develop their infrastructures in order to support instrument citation. PID providers need to ensure that efforts are coordinated, and not duplicated or working at cross purposes.

**What's next?**

# Workshop #2: Tallahassee/FSU

- August 20 - 22 at High Field Magnetic Laboratory
- Goals: Expand on workshop #1 by exploring pathways for PID implementation
- [Let us know if you want to attend!](#)
- We can support travel for some folks not at FSU.

# What's next?

- PID comparison
  - Syntax, governance, cost, process, metadata, persistence, etc.
- Focus groups
  - Facility managers, instrument operators, campus administrators, grant organization staff, and more
- Walking the walk
  - Ensuring that the facilities and instruments at our institutions have PIDs and encouraging their use
  - Looking into adding PIDs to institutional repository metadata and README templates

# Contact us / More Info

Project website:

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

Slack:

<https://fairosrcn.slack.com/>

# Jailbreaking toasters

- [Unauthorized Bread](#) by Cory Doctorow