FAIR Facilities and Instruments:

Enabling transparency, reproducibility, and equity through persistent identifiers

NSF FAIR Open Science (FAIROS)

Research Coordination Network (RCN)

NSF Awards #2226396, 2226397, 222639

Presented by Claudius Mundoma, ABRF 2025 NERLSCD Chapter

Organizations & Personnel

Key Questions for the FAIR-ROS Project

What are your main reasons for assigning PIDs to facilities and/or instruments?

What questions need to be answered?

What guidance is needed?

What outcomes and products from the project would be most useful for researchers?

Project Goals

Develop	Compile	Facilitate	Produce
Develop a Research Coordination Network (RCN) focused on the assignment of Persistent Identifiers (PIDs) to research facilities and instrumentation	Compile use cases for why and how PIDs might be assigned to facilities and instruments	Facilitate the generation of expertise and guidance on the key topics of interest	Produce & refine recommendations and lessons learned targeted toward the specific use cases

Key Questions

FAIR PRINCIPLES

Findable (F)
How do we enable people to find relevant facilities or instruments?

Accessibility (A)
How do we enable facilities and instruments to be accessible by wider audiences?

Interoperability (I)
How do we consistently capture relationships between persistent identifiers?

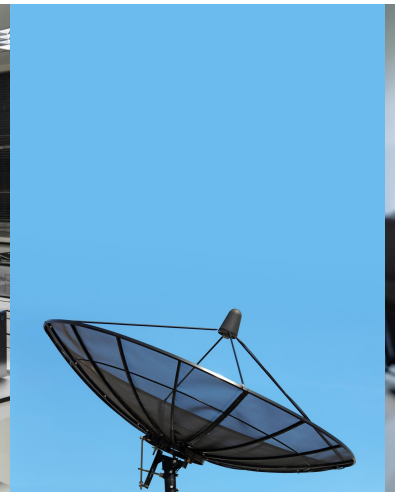
Reusability (R)
How can we incorporate information about facilities and instruments into data set provenance metadata more consistently?

Note on Terminology

Definitions are important - but are not consistent

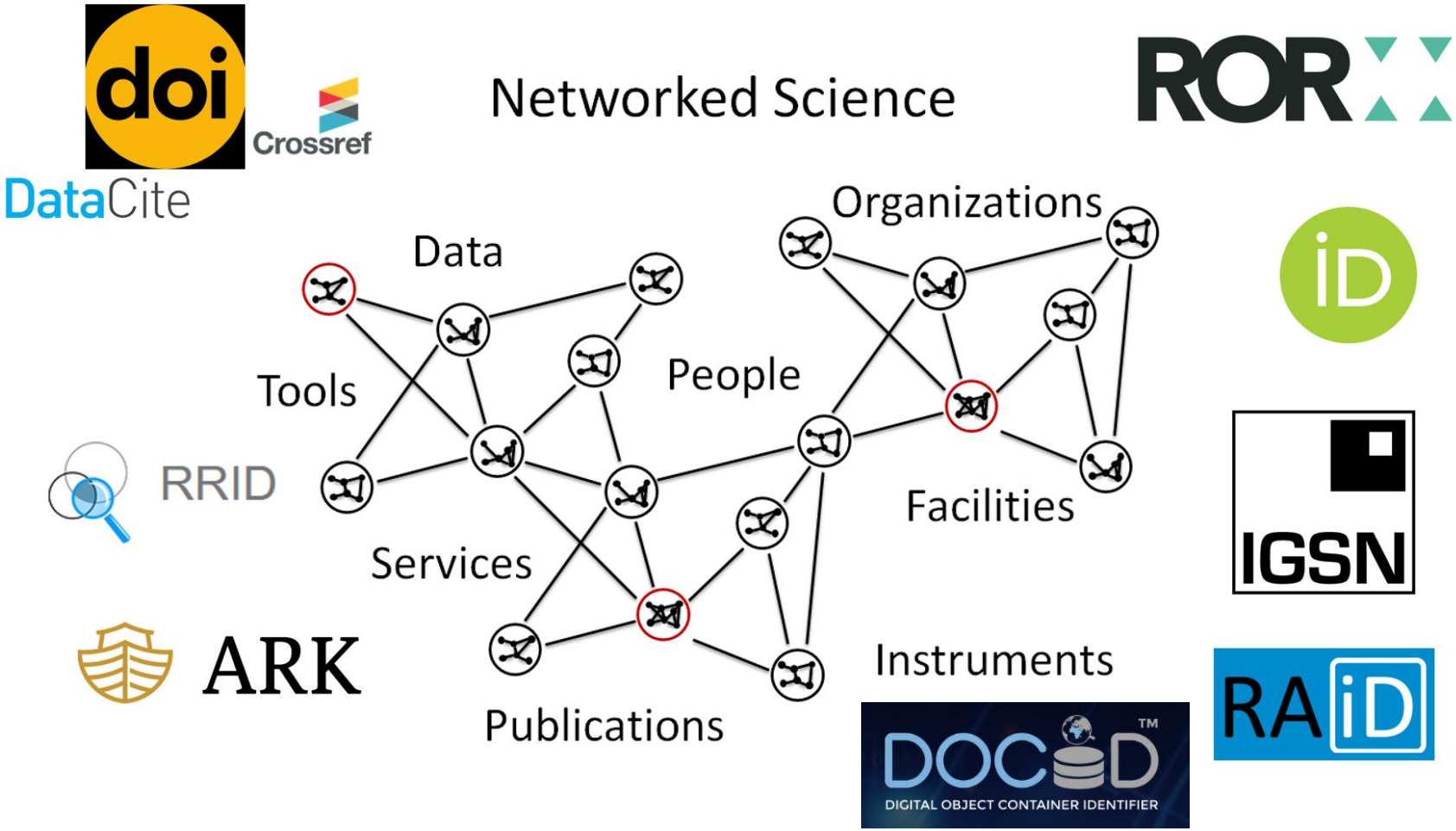
a. What is a “facility”?
b. What is an “instrument”?
c. What other terms are used? (platform, site, core, device, …)
d. When does it matter?





AT START: Less consistent Application of PIDs

Persistent IDs: Use cases



PIDs for Facilities and Instruments - NCAR

Example: doi Implementation

https://doi.org/10.5065/D6ZC80XJ

https://doi.org/10.5065/D6WM1BG0

UC Davis/NIH NeuroMab Facility

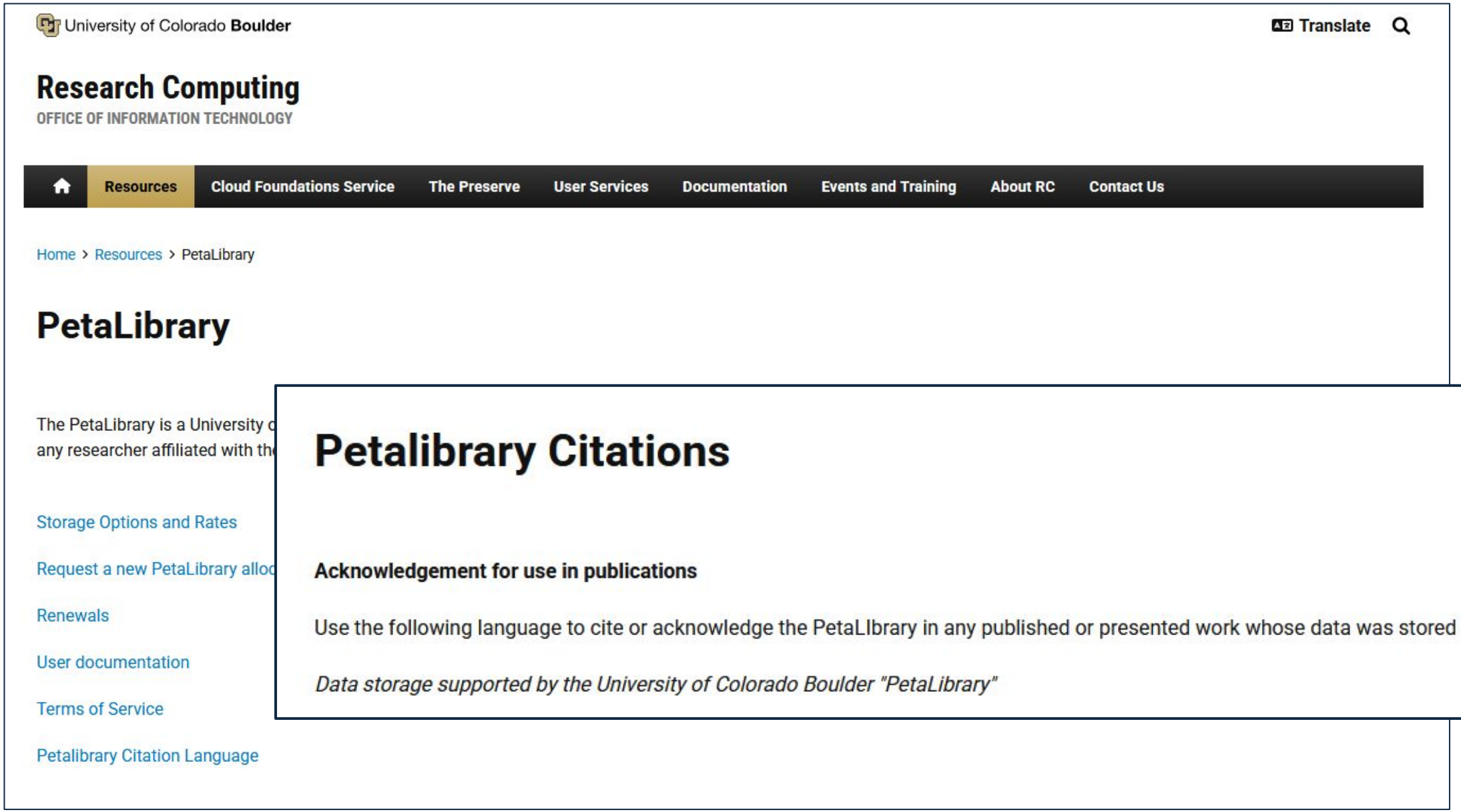
RRID

https://www.scicrunch.org/resolver/RRID:SCR_003086

https://neuromab.ucdavis.edu/

ROR

https://ror.org/00fyrp007



https://doi.org/10.25811/81nc-wv41

https://www.scicrunch.org/resolver/SCR_019299

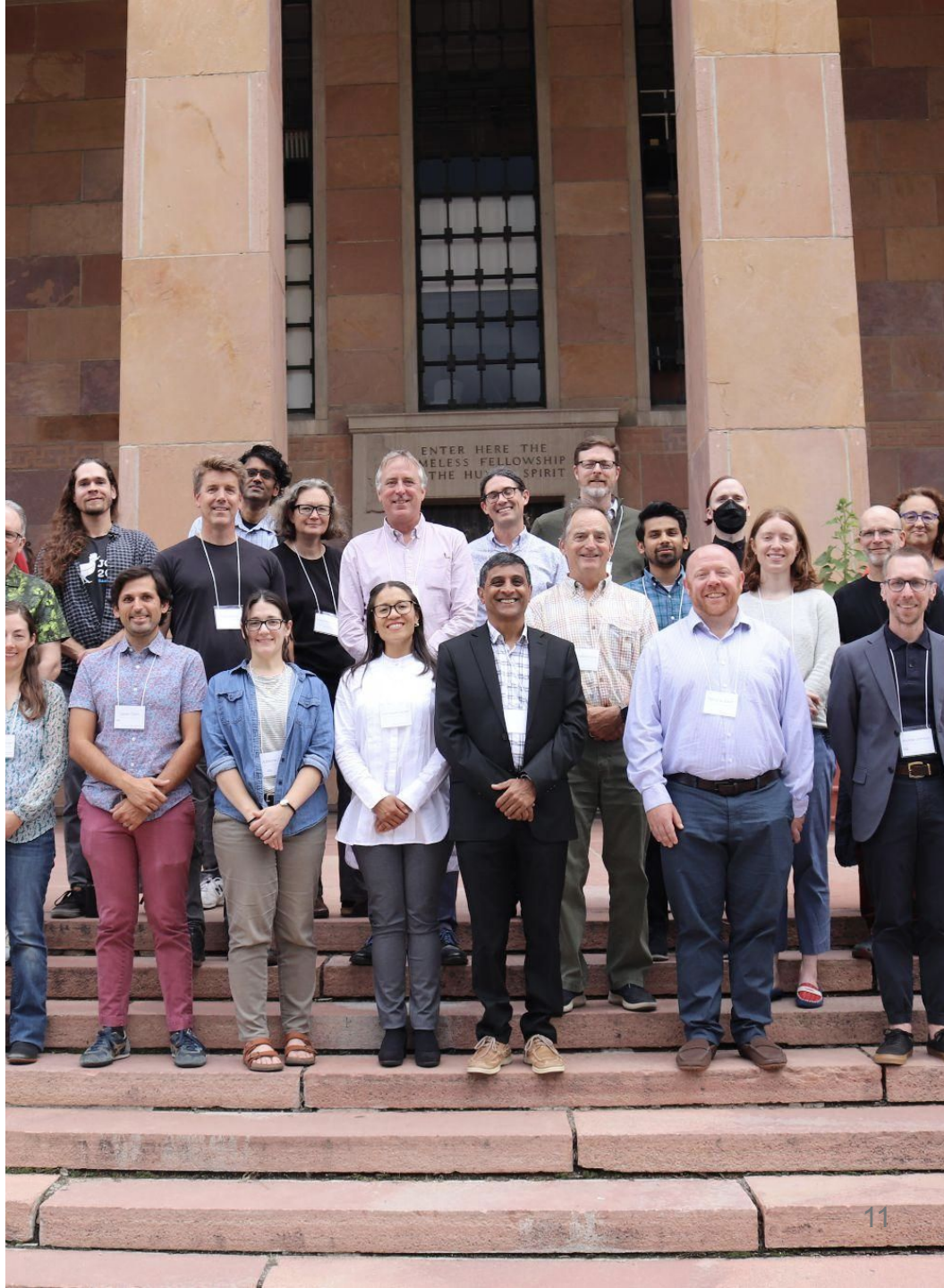
Project Advisory Committee

- **Anita Bandrowski** - founder and CEO of SciCrunch
- **David Butcher** - FAIR data management specialist at the National High Magnetic Field Laboratory
- **Matthew Buys** and **Kelly Stathis** - Executive Director and Technical Community Manager at DataCite
- **Zach Chandler** - Director of Open Scholarship Strategy, Stanford University
- **Nate Herzog** - CoreMarketPlace project lead at Vermont Genetics Network
- **Kevin Knudtson** - President of the Association of Biomolecular Resource Facilities (ABRF)
- **Giri Prakash** - Section Head of the Earth System Informatics and Data Discovery section at Oak Ridge National Laboratory
- **Dylan Ruediger** - Senior Analyst at Ithaka S+R
- **Shawna Sadler** - Head of Outreach & Partnerships at ORCID
- **Shelley Stall** - Sr. Director for Data Leadership at American Geophysical Union (AGU)

Workshop #1: September, 2023 – Boulder, CO

- **Need: PIDs are essential for scientific reproducibility, data provenance, and crediting instrument providers**
- **PID Systems: Current PID usage is scattered and inconsistent across different systems used for research instrumentation**
- Adoption: The focus should be on lowering adoption barriers and communicating value rather than choosing specific PID systems
- Metadata: Consider metadata alongside PIDs - PIDs alone cannot solve all challenges
- Granularity: Start simple with granularity and evolution tracking, then increase complexity only as needed
- Resources: Instrument/facility providers face significant resource limitations in assigning and managing
- **PIDs Value: Demonstrating clear value to users is critical for driving PID adoption and citation**
- Incentives: Different stakeholders (researchers vs administrators) require different incentives for PID adoption

Workshop #1 report: http://doi.org/10.5065/zgsx-2d06



Workshop #2: August 2024- Tallahassee, FL

Emerging topics

- **Need for facility and instrument PID recommendations as part of a national PID strategy**
- Need for more robust infrastructure and services for facility and instrument PIDs
- **Engagement needed with instrument manufacturers to adopt PID-supporting practices**
- Engagement needed with journal publishers and editors on PID incorporation

Workshop #2 report:
http://doi.org/10.5065/jea7-yf24

Workshop #3: August 2024- NCAR -Boulder, CO

Emerging topics

- **Developed recommendations for PID implementation**
- Formed consensus on which PID to use for what use case
- **New project to look into equipment PIDs - focus on granularity**
- Need to engage research enterprise software providers
  - Reference Managers
  - Stratocores
  - iLab
  - Protocols.io
  - etc…

Workshop #3 Draft Recommendations

PID Curators: August 2024- NCAR -Boulder, CO

ROR ORCID RRID

Common Themes

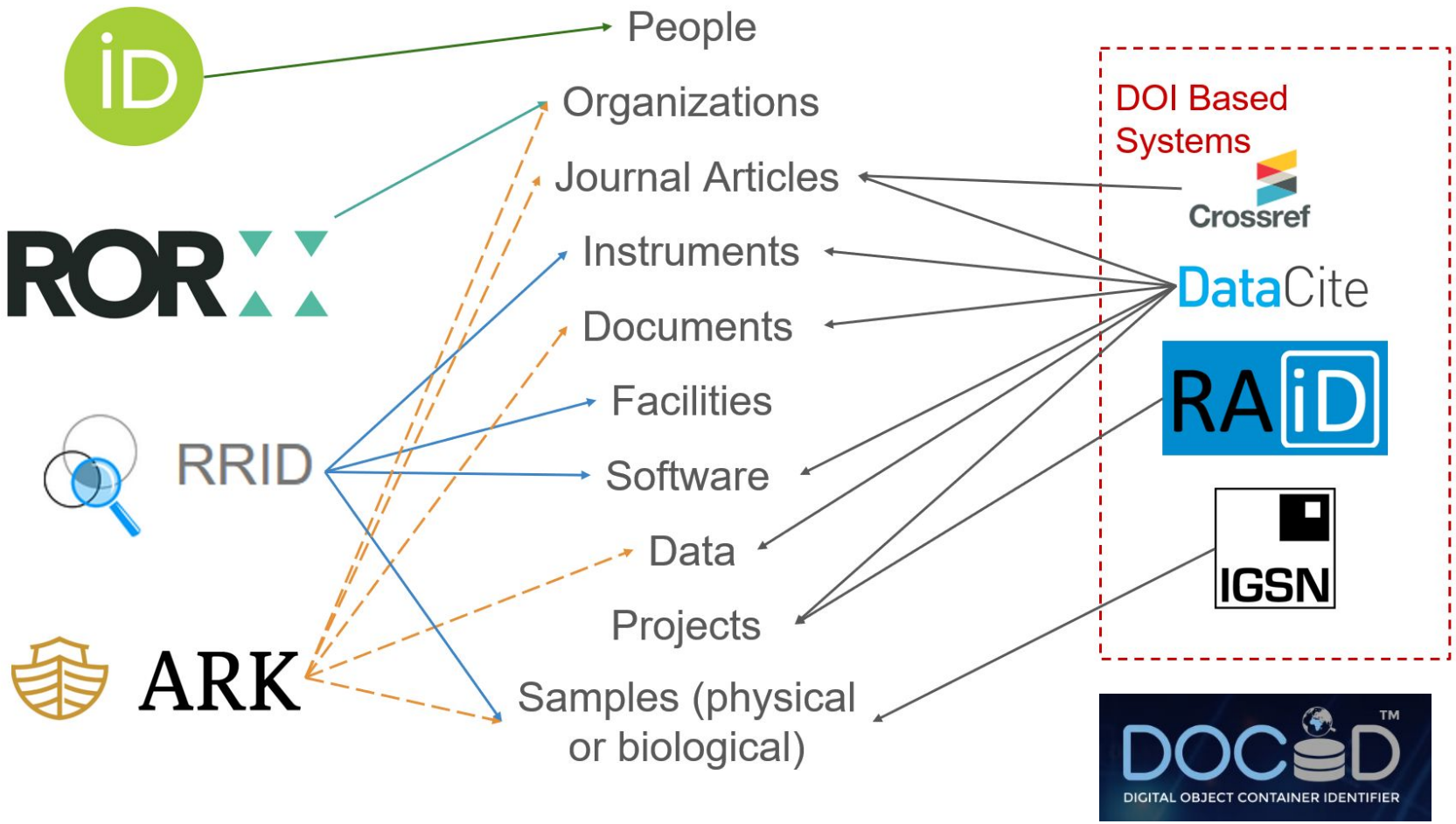
1. Use cases matter
    a. Reproducibility and Replicability
    b. Data provenance
    c. Attribution: Track impact and citations
    d. Discoverability and Collaboration: Find and share resources
2. PIDs are a starting point
    a. Making PIDs and citations visible and actionable for researchers who use facilities and instruments is critical
    b. Value from PIDs comes from integrating them into other systems (metadata systems, institutional systems, publishing systems)

Recurring questions

- What metadata needs to be included? Where should the metadata be collected and made available?
- 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 instrument or do you need a PID specifically for components?
- Scientific Instrument of Theseus
  - Instruments and facilities evolve over time
  - When is a new PID issued vs. metadata updated?
  - New software? New hardware?

PROGRESS: More consistent Application of PIDs

Persistent IDs: Use cases

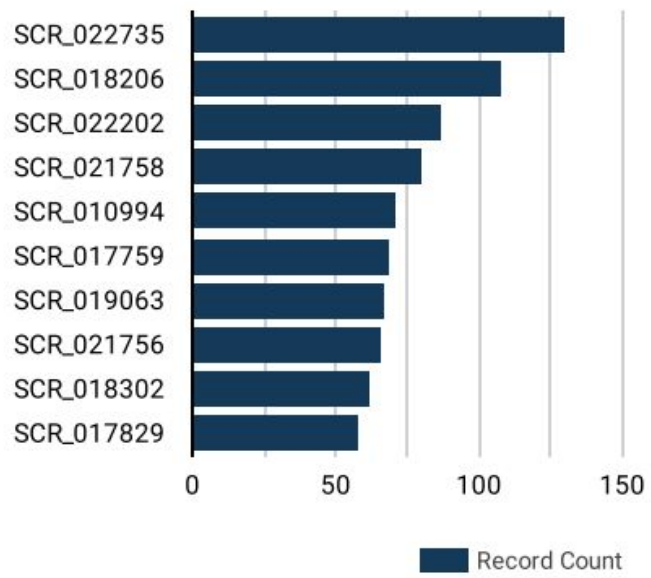


CORE FACILITIES REPORT

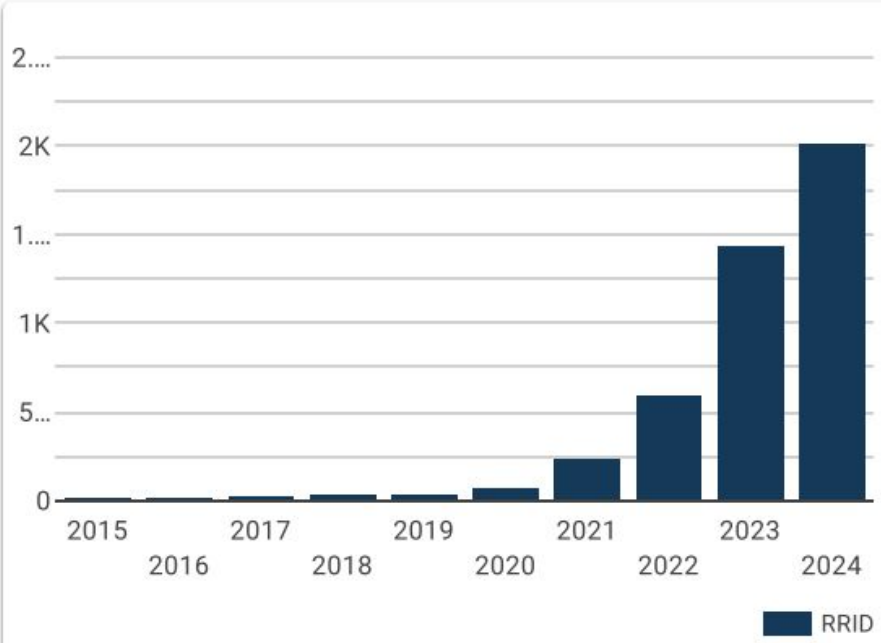
CORE FACILITY | PARENT ORGANIZATION | SOURCE OF DA... (1) | YEAR | JOURNAL | SNIPPET/GRANT NUMBER

CORE FACILITY	PMID	JOURNAL
552	3,816	815

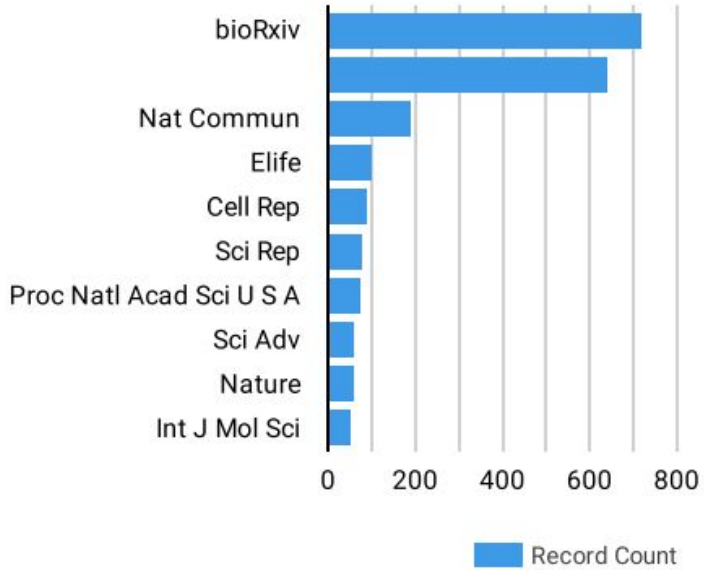
CORE FACILITIES



NUMBER OF PUBLICATIONS FOUND BY GRANT OR RRID

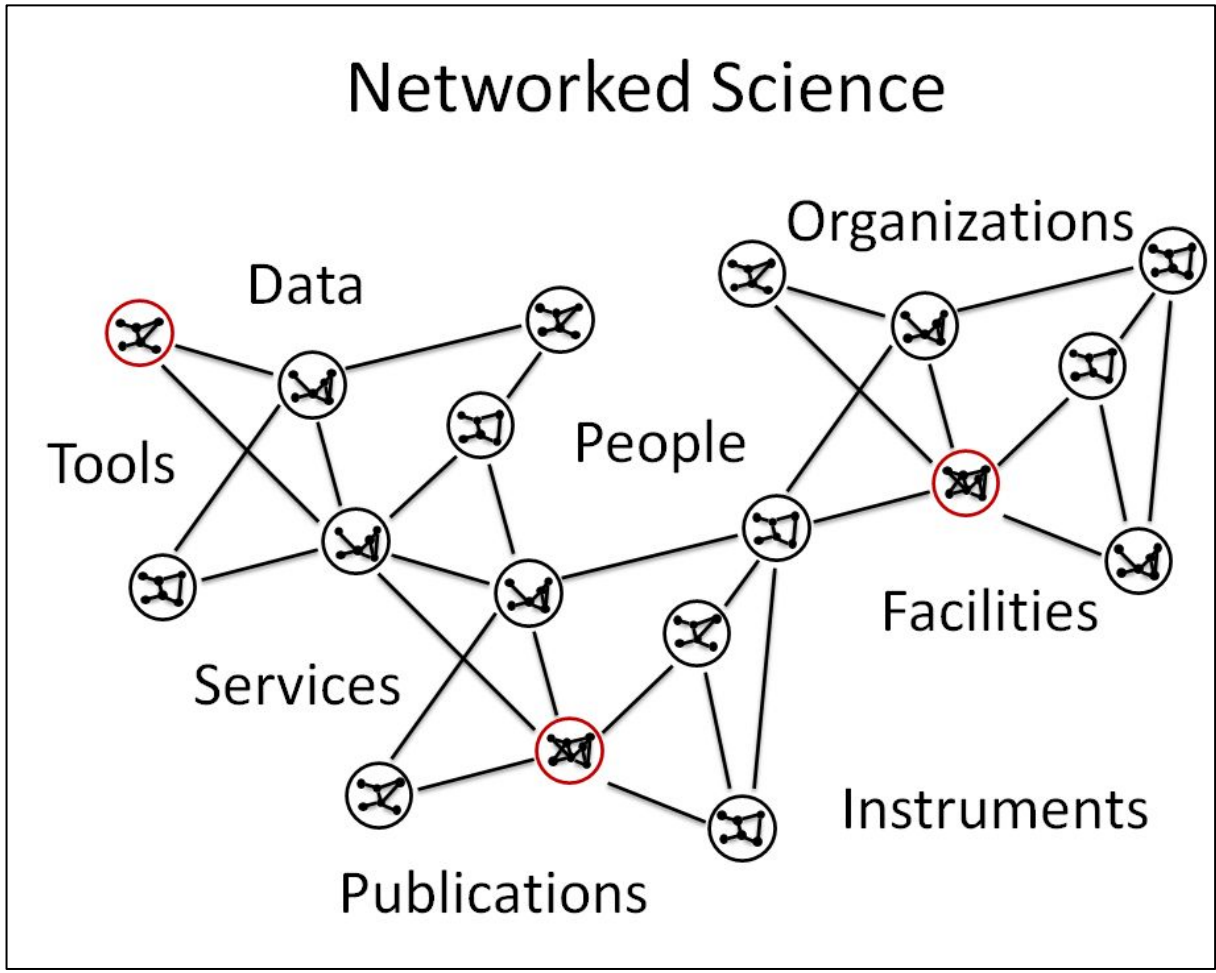


JOURNALS



Creating and Maintaining 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 are up to date?
- How are connections between PIDs to be created and maintained?
- Where is funding going to come from?



Multiple Stakeholders - Distributed Responsibilities

Academic research institutions

National laboratories

Nonprofit organizations

Instrument manufacturers

Facility and instrument operators

Research scientists/users

Publishers and editors

PID system providers (RRID, DOI, ROR)

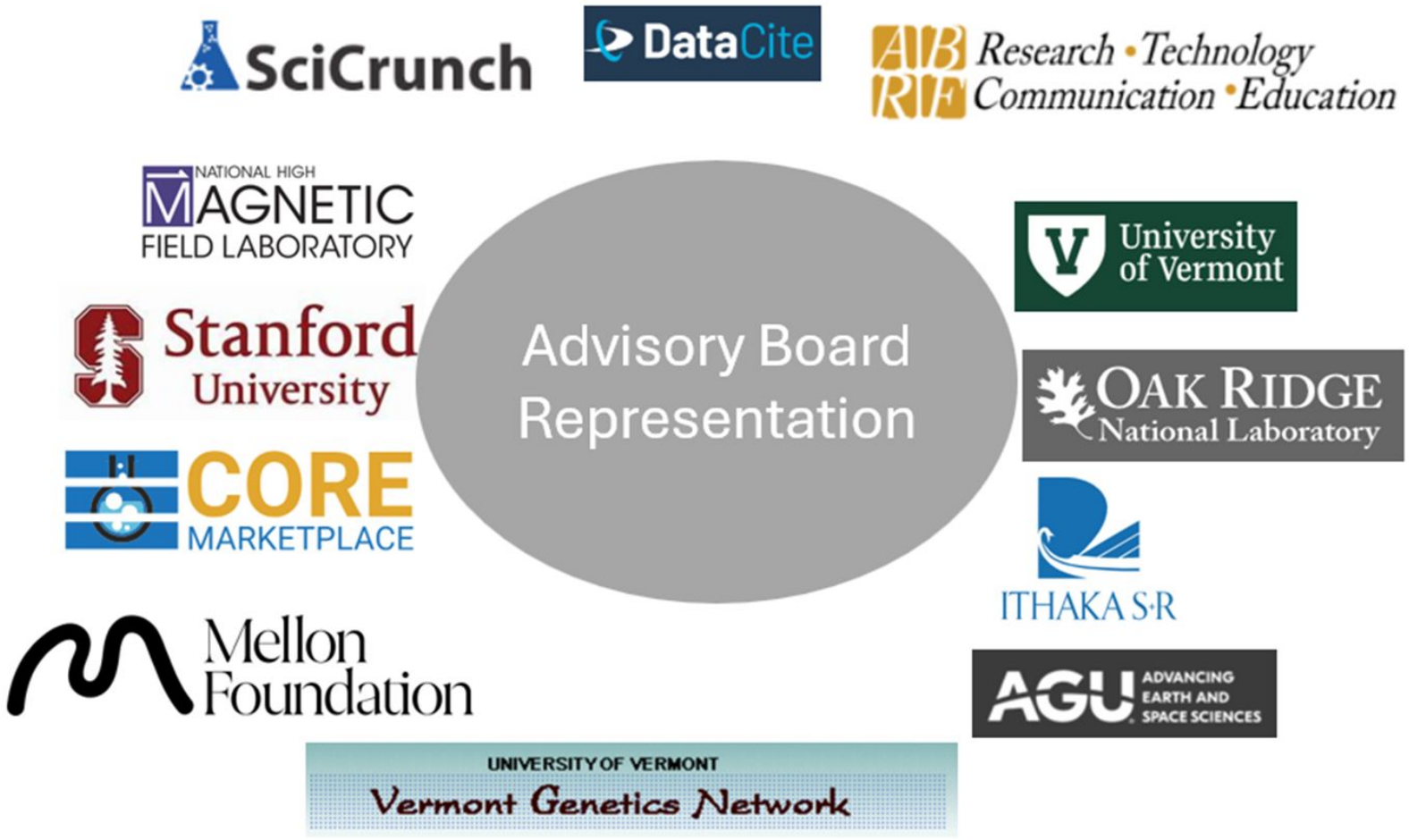
ACT!:

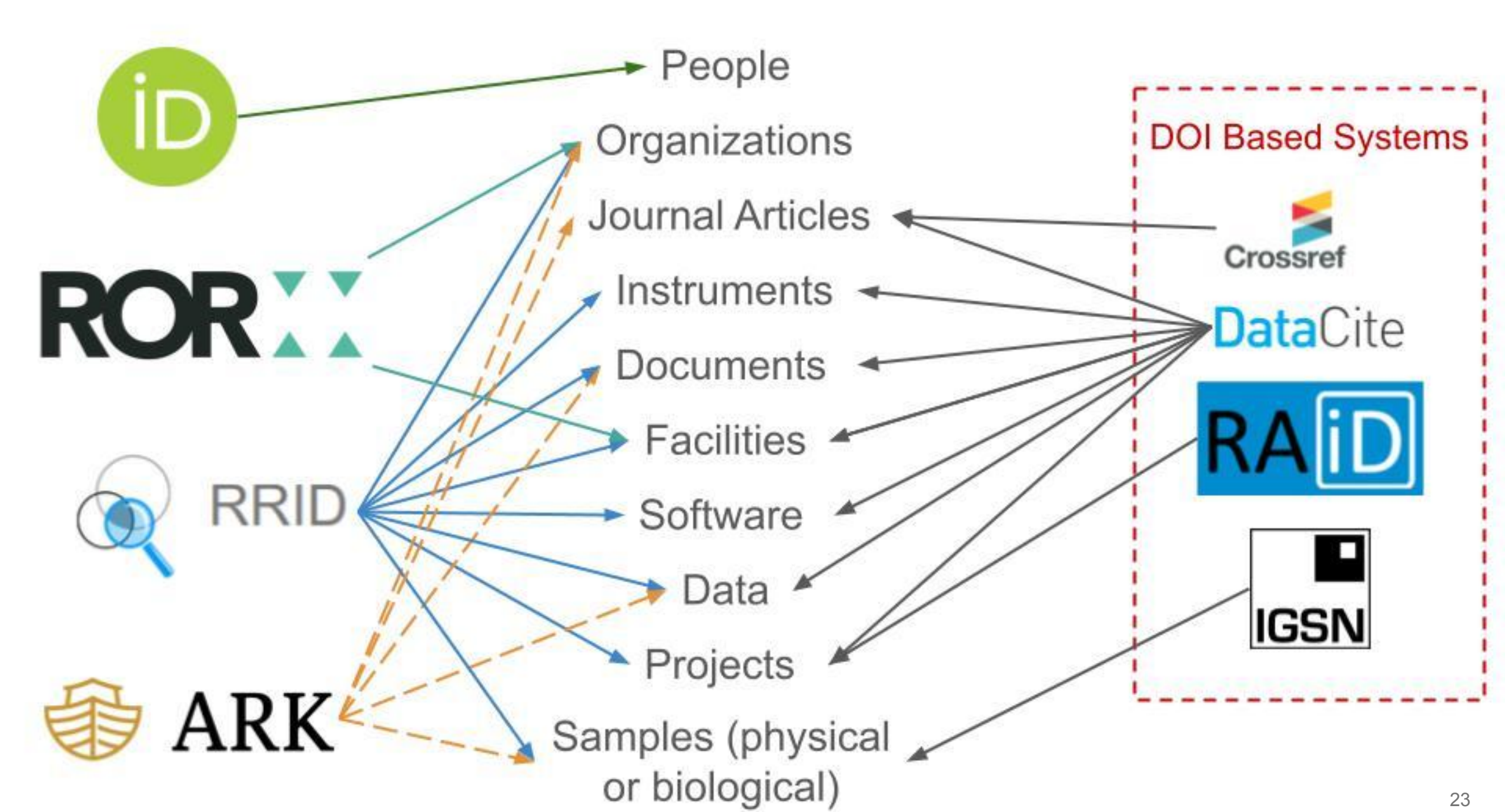
Persistent IDs: Call to ACTION

This is bigger than Rigor & Reproducibility
- Its Securing Funding,
- Core Facility workflows,
- Connected Labs
- Proper acknowledgement

PID MVP

THANK YOU!


SciCrunch
DataCite
AIB RIF Research • Technology Communication • Education
NATIONAL HIGH MAGNETIC FIELD LABORATORY
Stanford University
CORE MARKETPLACE
Advisory Board Representation
University of Vermont
OAK RIDGE National Laboratory
ITHAKA S+R
Mellon Foundation
AGU ADVANCING EARTH AND SPACE SCIENCES
UNIVERSITY OF VERMONT
Vermont Genetics Network


People
Organizations
Journal Articles
Instruments
Documents
Facilities
Software
Data
Projects
Samples (physical or biological)
DOI Based Systems
Crossref
DataCite
RAiD
IGSN
ROR
RRID
ARK

Motivating Examples

Connecting Facilities, Instruments, & Data

NSF NCAR HIAPER Gulfstream GV

https://doi.org/10.5065/D6DR2SJP

ACCLIP NSF/NCAR GV Instrument Data Merges - 10 Second
https://doi.org/10.26023/2HAX-YPQB-GG0Q

FAIRO-1 Ozone Data
https://doi.org/10.26023/S3FA-R52G-ZS11

HIAPER Atmospheric Radiation Package (HARP) CCD Actinic Flux Spectrometers Photolysis Frequencies
https://doi.org/10.5065/D6MP51N7

…

[686 datasets]

Project Activities

- Focus groups & presentations to relevant groups
  - NSF FAIR Open Science RCN project cohort
  - Earth science facility providers and users
  - FSU & CU campus facilities staff
  - CI Compass - FAIR Data Working Group
  - Data Curation Network
- Conference engagement - AMS, ABRF, ESIP, IASSIST, RDA, RDAP, Year of Open Science
- Sept 2023 Boulder Workshop - 35 participants
- Aug 2024 Tallahassee Workshop - 35 participants