

# There will be FAIR. Insights into the PID landscape analysis

PID4NFDI - Persistent Identifier Services for NFDI

Torsten Kahlert, Sara El-Gebali, Stephanie Hagemann-Wilholt *on behalf of the PID4NFDI Team*  
User Conference 4 Base4NFDI | 19. – 20.11.2024 | Berlin

Funded by  
 Deutsche  
Forschungsgemeinschaft  
German Research Foundation

Funded by DFG as part of NFDI.  
Grant Number: 521466146

# Why PIDs matter

## *The Potential of PIDs*

- persistent, globally unique and resolvable **identifier** attached to digital resources (e.g., publication, dataset)
- key for FAIR RDM: **identifies & link** research assets
  - enhances data findability, accessibility, citability
- accompanying metadata
  - standardized, interoperable
  - rich and sustainable descriptions
  - incorporation of other PIDs for resource connections
  - reduces admin work & errors through automated processes and validation
- useful for the whole research lifecycle



# No PID? Not FAIR!

To be Findable:

- **F1. (meta)data are assigned a globally unique and eternally persistent identifier.**
- **F2. data are described with rich metadata.**
- **F3. (meta)data are registered or indexed in a searchable resource.**
- **F4. metadata specify the data identifier.**

To be Accessible:

- **A1 (meta)data are retrievable by their identifier using a standardized communications protocol.**
- A1.1 the protocol is open, free, and universally implementable.
- A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
- **A2 metadata are accessible, even when the data are no longer available.**

To be Interoperable:

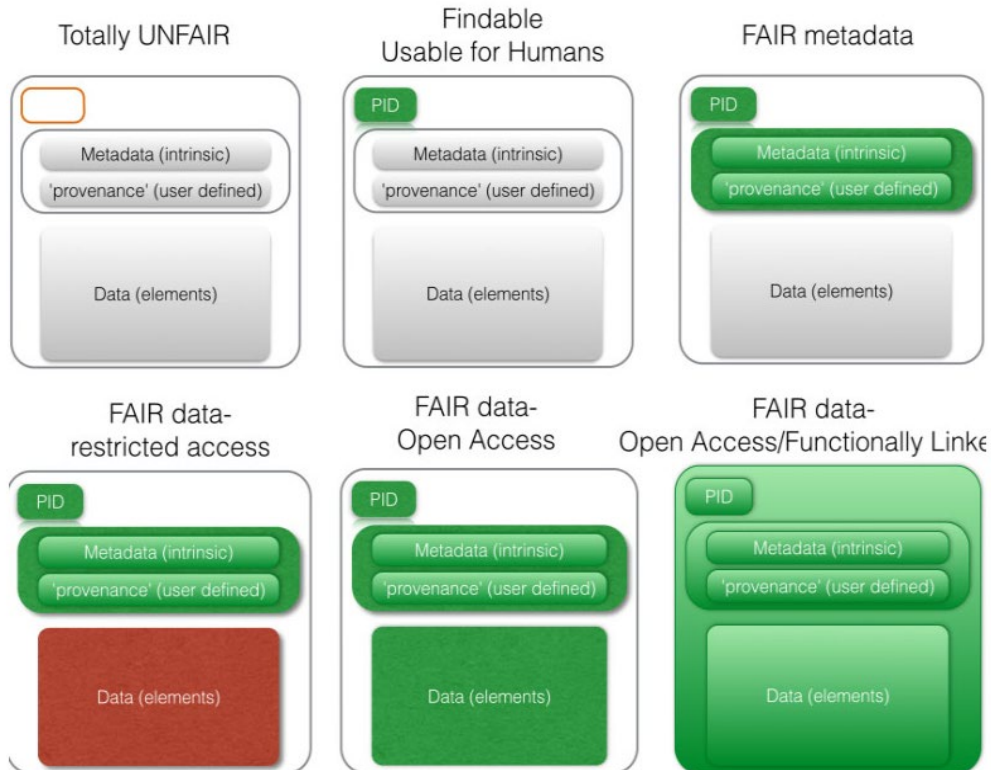
- **I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.**
- I2. (meta)data use vocabularies that follow FAIR principles.
- **I3. (meta)data include qualified references to other (meta)data.**

To be Re-usable:

- **R1. meta(data) have a plurality of accurate and relevant attributes.**
- **R1.1. (meta)data are released with a clear and accessible data usage license.**
- **R1.2. (meta)data are associated with their provenance.**
- **R1.3. (meta)data meet domain-relevant community standards.**

# No PID? Not FAIR!

## Data as increasingly FAIR Digital Objects



# Explain the service

*Service and technical concept*

**Objective:** *evolve PID service landscape within NFDI at all levels (technical, organisational, methodological, and in communication)*

- integrate existing PID infrastructures
- improve metadata quality and mapping
- training & outreach
- support PID integration in consortia
- align with national PID strategy (PID Network)

Initialisation

Integration

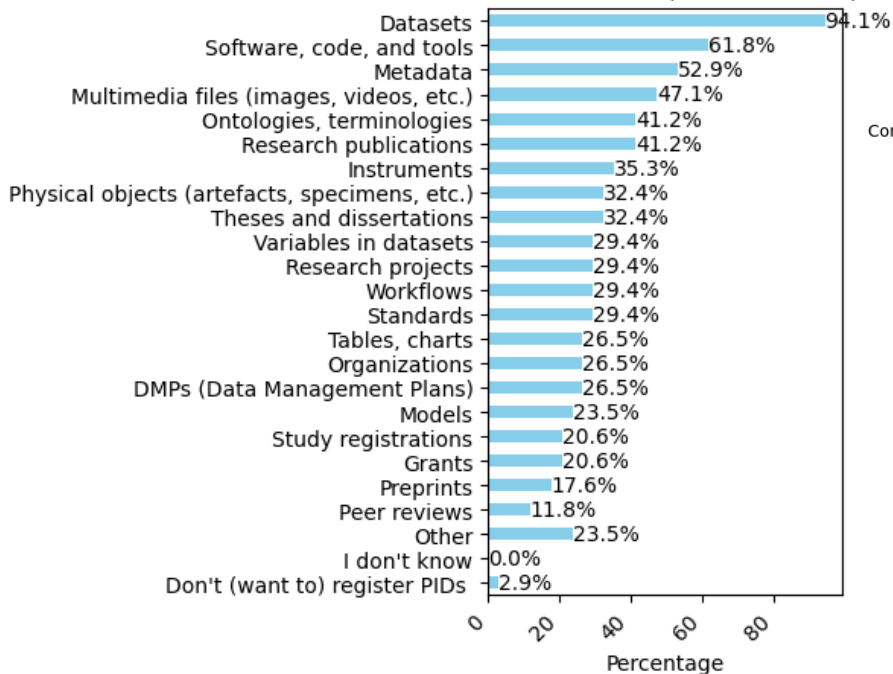
Ramp-Up

# Preliminary Survey Results

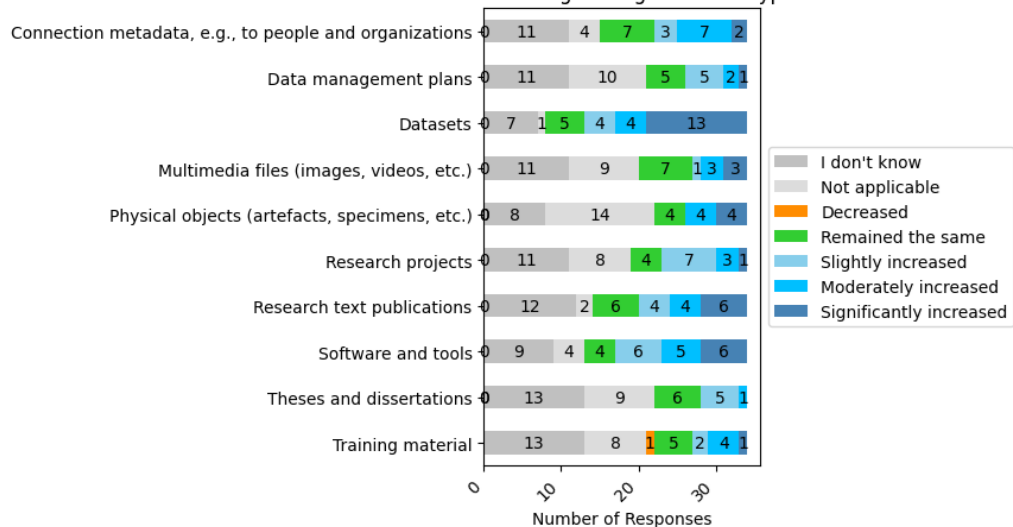
## Landscape Analysis

# Resource Types & Trends

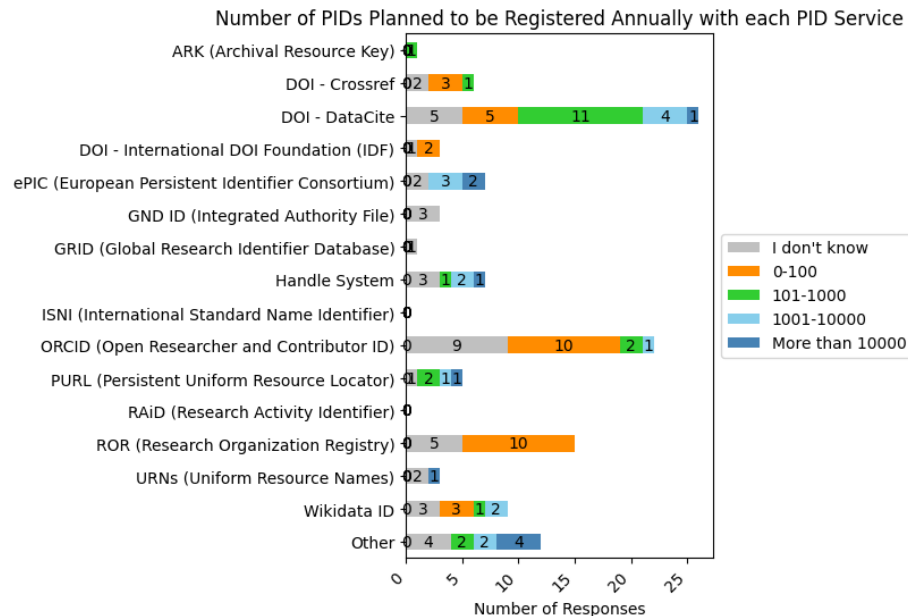
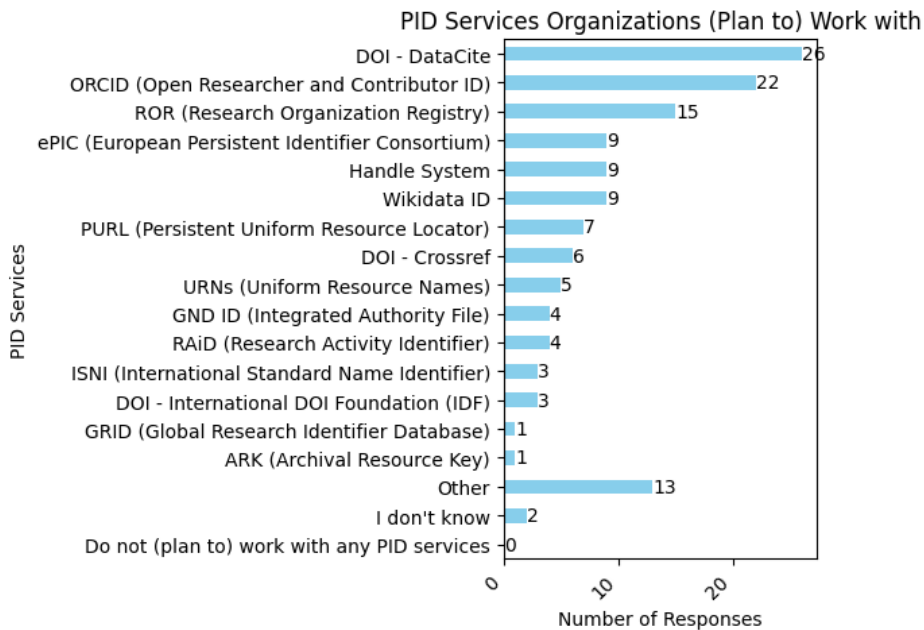
Entities for which PIDs are (Planned to be) Registered



Trends in Registering Resource Types



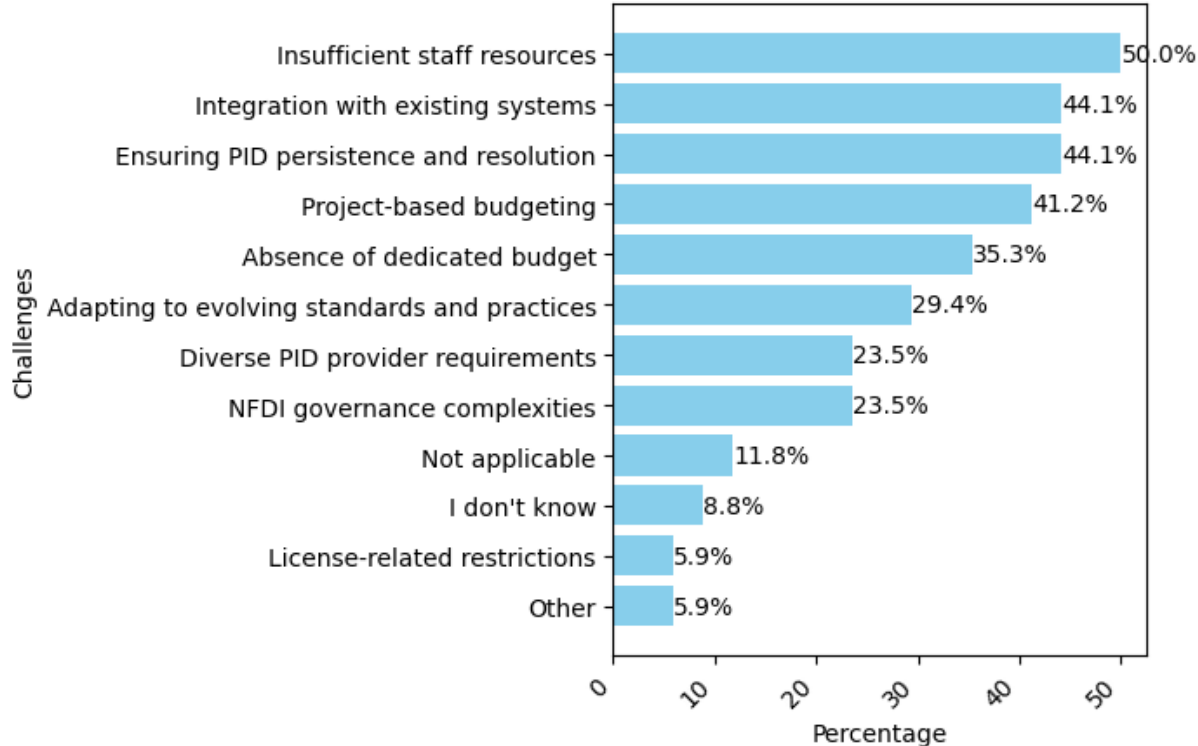
# PID Provider Affiliations & Scaling





# Challenges in PID Governance

Challenges within Organization regarding the Governance of PID Implementation and Management



# The PID Coordination Hub

# Vision: PID Coordination Hub

## Governance

- Decision matrix for PID (provider) selection
- NFDI-wide PID policy
- Compliance Testing with existing PID policies

## Metadata & Interoperability

*based on high TRL services*

- Tools and support for metadata quality assessment and monitoring
- Development of guidelines and best practice recommendations for metadata harmonization
- Data Type Registry
- PID Metaresolver

→ create foundation for NFDI  
PID Graph

## Support & Information

- Linking existing PID resources and services
- Display best practice according to use cases
- Helpdesk as first level support
- Training sessions & consultations

## Use cases

*special focus on:*

- Instruments
- Physical Objects
- High-granular/ warm data
- Projects & grants (explorative)

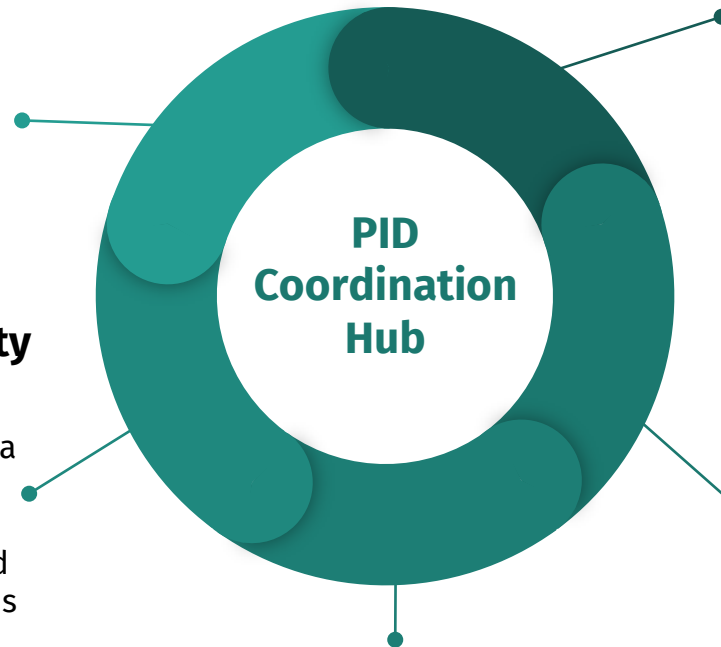
Focus  
Groups

## Target groups

*community engagement with:*

- Repository managers
- Researchers
- Research organizations
- Decision makers

→ representatives of NFDI bodies



# PID Provider Selection Guide

- Overview of PID providers for NFDI use cases
- Support for selection of PIDs and PID providers (considering POSI principles)
- Transparency of selection criteria (e.g. metadata policies)

Persistent Identifier Guide

dutch digital heritage network

Introduction Theme 1 Theme 2 Theme 3 Theme 4 Theme 5 Result

Aims (What do you want to achieve by implementing Persistent Identifiers?)

Position 1

I want a PID system that comes with its own metadata policies and requirements.  
*Some PID systems come with metadata policies and requirements, e.g. regarding required metadata fields.*

Strongly Disagree Disagree Neutral Agree Strongly Agree

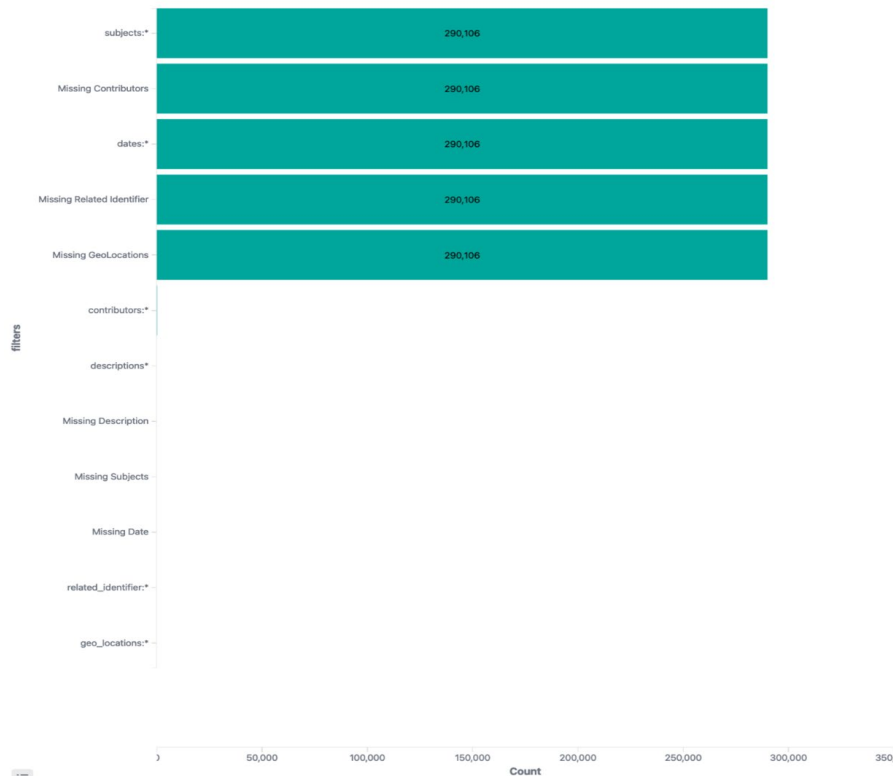
important

Level of preference for Persistent Identifier solution

Provider	Preference Level (0-100)
Handle	40
URN:NBN	80
DataCite DOI	80
ARK	40

The more freedom you require in terms of policy creation, the more internal regulation you will have to govern and maintain. URN:NBNs and DataCite DOIs have clear (and strict) policies, while choosing ARKs and Handles means you are quite free to create your own policies and apply them as strictly as you want.

# Metadata completeness



## Missing (sub-)properties in **Recommended** fields:

- related Identifier → lack of related resources means it is not linking the resource to other outputs
- geolocation
- contributors

# Training Concept

Modular approach for practical guidance

Target groups: infrastructure managers, data managers/curators, researchers, trainers/intermediaries, decision makers

Level 1 – Fundamental

- Explaining benefits & providing overview

Level 2 – Advanced

- targeted training materials, integration of multiple PIDs, best practices

Level 3 – Expert

- covering research (data) lifecycle with connected PIDs, customized support

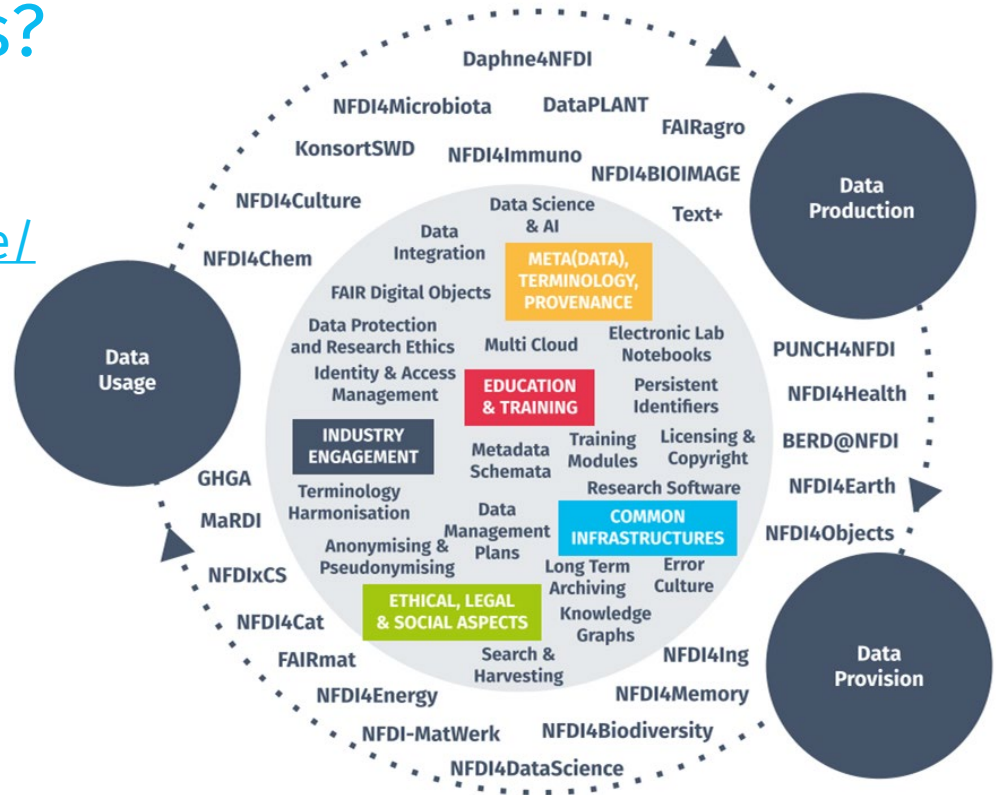
# Thank You! Questions?

More information:

<https://pid.services.base4nfdi.de/>

Contact us:

pid4nfdi@lists.nfdi.de



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