Pure Support for Research Ecosystems through National Persistence Identifier Strategies

Initial Project definition document for discussion

Malene Knudsen, May 2025

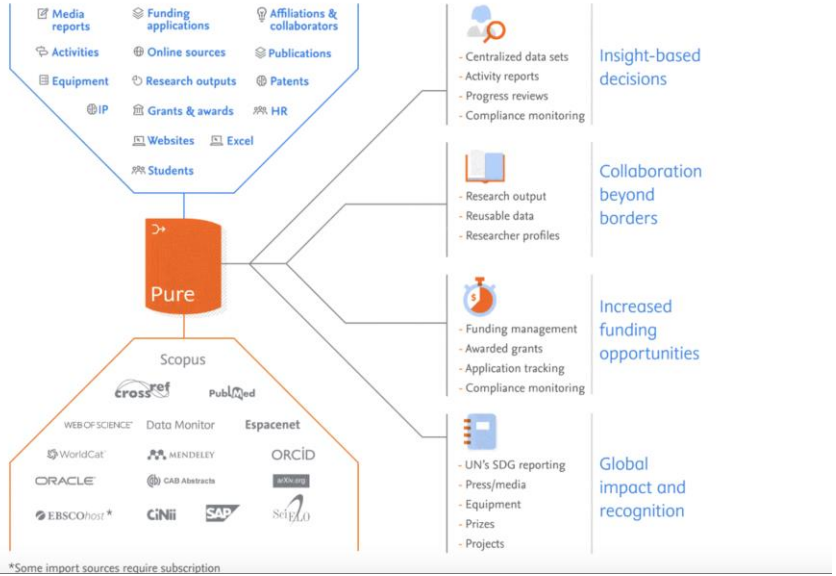
Introduction

To Me:

- Senior Product Manager of Integrations
- 24 Year of Pure experience
  - 10 years as a client
  - 14 years with Elsevier
- Librarian
- Mother of 3

To Pure:

- CRIS system
- 23 years old
- 550+ clients all over the world



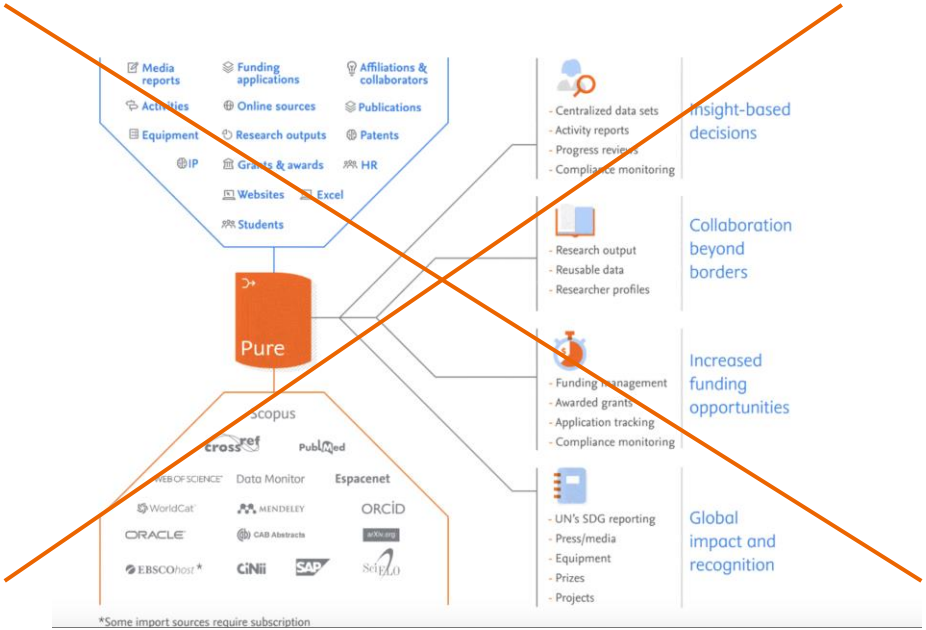
Introduction

To Me:

- Senior Product Manager of Integrations
- 24 Year of Pure experience
  - 10 years as a client
  - 14 years with Elsevier
- Librarian
- Mother of 3

To Pure:

- CRIS system
- 23 years old
- 550+ clients all over the world



Introduction and Context

Rise of National PID strategies

What are National PID Strategies?

- National Persistent Identifier (PID) strategies are coordinated efforts by countries to implement a framework for the use of PIDs across their research ecosystems. They are core to the concept of FAIR* data principles and aim to standardise the identification of research outputs, such as datasets, publications, project, and researchers, facilitating better data management and accessibility

- **Key Components of strategies:**
  - **Interoperability:** Ensures that different systems can work together seamlessly.
  - **Standardisation**: Establishes common practices for PID usage across various platforms
  - **Governance:** Involves creating frameworks for managing and supporting PID systems.

* Wilkinson, M.D. *et al.* (2016) “Comment: The FAIR Guiding Principles for scientific data management and stewardship,” *Scientific Data*, 3. Available at: https://doi.org/10.1038/sdata.2016.18.

Why have national PID strategies become important?

- **Enhanced Discovery and Accessibility:** PIDs provide reliable methods for accessing research outputs, improving data quality and making it easier for researchers and the public to find and utilise data.
- **Reduction of Administrative Burden:** By streamlining processes related to managing research outputs, PIDs help reduce time and costs associated with administrative tasks.
- **Improved Research Assessment:** PIDs enable better evaluation of research impact through consistent metrics, which is crucial for funding and policy decisions.
- **Transparency and Accountability:** They foster trust within the research community by enhancing the visibility of research activities and outputs.
- **Facilitation of Global Collaboration:** PIDs support international research efforts by ensuring that different systems can communicate effectively, promoting cross-border partnerships.
- **Support for Open Science Initiatives:** National PID strategies align with broader goals of open science, making research more accessible and reusable.

Example National PID strategies – so far (more in early development)

Country	Strategy Development Stage	Implementation Level
Australia	development	not started
Canada	development	partially implemented
Czech Republic	implementation	partially implemented
Finland	development	partially implemented
Germany	development	partially implemented
Netherlands	implementation	partially implemented
New Zealand	implementation	partially implemented
South Korea	implementation	partially implemented
United Kingdom	implementation	partially implemented
Ireland	Planning (recommendations made)	not started
United States	Planning (recommendations made)	not started

1. https://www.pid-network.de/en/news/blog/use-cases-nationalen-pid-strategien
2. https://upstream.force11.org/desirable-characteristics-for-persistent-identifiers/
3. https://www.niso.org/niso-io/2024/08/national-pid-strategies
4. https://www.crkn-rcdr.ca/en/developing-canadian-pid-strategy-results-and-next-steps
5. https://wiki.surfnet.nl/display/DRIVERguidelines/Annex+Use+of+Persistent+Identifiers
6. https://www.rd-alliance.org/groups/national-pid-strategies-wg/outputs/?output=94446
7. https://www.jisc.ac.uk/innovation/projects/a-national-persistent-identifier-research-strategy
8. https://www.clarin.eu/content/persistent-identifiers
9. https://pidforum.org/t/developing-a-us-national-pid-strategy-report/2839
10. https://libguides.utsa.edu/federal_policies/PID
11. https://pro.europeana.eu/index.php/post/policy-for-persistent-identifiers-in-the-data-space
12. https://repository.dri.ie/catalog/sn00qt29n

Noted challenges in adopting national PID strategy

- **Problems groups are trying to solve**
  - Fragmentation and lack of coordination in the PID landscape
  - Risk of over-reliance on PIDs in the ecosystem without ensuring long-term sustainability (commercial vs community-led)
  - Lack of awareness and adoption of PIDs among some stakeholder groups (i.e. users/funders of research + CRIS users)
  - Potential technical failures or discontinuation of PID services (see sustainability issue)
  - PIDs not being actionable

Pure PID response

Proposed project activity and communications

Value Proposition for implementing a PID strategy in Pure

A well-functioning PID infrastructure connects all elements of the research lifecycle. PIDs are essential for modern research and associated research systems, ensuring unique identification and accessibility of research components. They contribute to research integrity, reproducibility, discoverability, and efficient information management

- **Opportunities**:
  - Enhanced research connectivity (data is linked and reusable)
  - Improved automation and efficiency (time saving for admin/researchers)
  - Better research monitoring and assessment (for research information management)
  - Improved open research infrastructure (aligning to FAIR principles)

Discovery underway

1. Understand recommendations being made across global Pure communities
2. Review/Audit of existing PID support and provision in Pure
3. Scope of possible PID project, including priority requirements and business case for resource investment
4. Roadmap implications and cross-product requirements

Community engagement survey

Survey for Pure community

What are considered as the priority PID and workflows

How are PID strategies being communicated within your institutions?

Are there any quick wins that can help the community?

How can we build a collaborative approach to metadata in support of PID across Pure UG?

Can we Improve knowledge of what is possible in Pure already?

Are there opportunities for workflow reviews based on existing provision (i.e. ORCID integration)

Are there timing constraints or other dependencies (such as assessment exercises etc)

Result of the survey

Come back next year, and I will tell you the result ☺

Existing PID provision/support in Pure

ORCID

- Pure have integrated with Orcid since 2016
- ORCID Certified Partner as of 2025

- Import/Export via ORCID for authenticated researchers
- Use of authenticated ORCID in Import Sources
- Use of authenticated ORCID in Cronjobs
- Export of authenticated ORCID via OAI-PMH
- Use of ORCID to locate Scopus Author ID
- Showcase authenticated vs. Non-authenticated ORCID on Pure Portal (in development)

Research Organization Registry (ROR)

- On Internal organization
  - Used for more precise mapping of affiliation data
  - DataCite integration
  - Orcid integration

- External organisations
  - Used for more precise mapping of affiliation data

Digital Object Identifiers

DOI

Minting of DOI available via DataCite for:

- Dataset records (2014)
- Publications (2024)

Extensive use of DOI for search/discovery

Used of handling duplicates

Use of DOI to support OA workflows

Handle.net server

UUID based handles across all content

URN Support

URN support in some countries (i.e. NL)

Elsevier Identifiers

Extensive use of

- Scopus Author ID
- Scival Organisation ID
  - Import Scival Organisations
- Scopus Publication Identifier (EID)
  - Match publications with Scopus EID
  - Citation import (Scopus/Scival)
- Journal Info
  - ISSN
  - Scopus Journal ID

Pure development plans

High level scope of PID Integration work required

Project Initiation

- Requires cross team/product coordination and business case

Define scope and requirements

- Establish priority PIDs, such as ORCID, DOI, ROR, and RAiD.
- Define priority workflows to support – where can PIDs offer most help

UI and Data Model alignment

- Add PID fields to relevant content types and forms.
- Align Pure metadata schema with PID metadata schema (for priority choices).

Integrations

- Ensure the ability to create, read, update, delete PID records in support of workflow use cases
- Update all OAI-PMH formats to include PIDs
- Implement automatic metadata harvesting from PID providers
- Implement PID validation and lookup functionality
- Make PIDs actionable

Reporting

- Update reporting to accommodate new PID values

API/Sync framework

- Develop API and Sync framework to allow interaction with PID registries

High level scope of PID Integration work required (2)

GTM

- Communicate with user communities on existing levels of PID support
- Publicise value story of Pure PID support and use cases in territories with national PIS strategy adoption (aka EuroCRIS/Pure Conf)
- For example URN/ORCID/DOI case studies

Training/Documentation

- Update documentation regarding PID/PID usage
- Train client services/sales regarding Pure PID support

Risks

- Alignment with national PID strategy roadmaps (timing)
  - Many still work in progress
- Moving too early – what can be done now to support the requirements for national roadmaps?
  - What are the deadlines?
- Will the landscape/requiment change before we get to it
- ….

Thank you

Pure, Elsevier

Malene Knudsen, m.knudsen@elsevier.com