The Role of Universities in the Implementation of PIDs
The PID Graph

- Most recent (Aug 2020) detailed PID Graph snapshot at hand...

- ... but things have evolved quite a bit ever since: RORs/Ringgold IDs, DOI-based grant IDs, PIDINSTs, IGSNs, RAiDs...

- Some feeling of “building the plane as we fly it” – is it possible to keep track of this evolving snapshot?

*Figure: PID graph as described by FREYA in Aug 2020 – with a strong focus on research outputs*
The “Building the Plane” report – and associated case studies

- “... to identify, through investigation, analysis and recommendations, the best possible **strategic and operational paths** to achieve a well-functioning PID infra for KE member states and beyond”

- “… to identify the **main risks** when pursuing a well-functioning PID infrastructure for research, and to better understand the most important elements of **trust** in creating said infrastructure”

https://zenodo.org/record/7258286
The “Building the Plane” report – and associated case studies

Report:
Building the plane as we fly it: the promise of Persistent Identifiers, [https://zenodo.org/record/7258286](https://zenodo.org/record/7258286)

Case studies:

- Adoption of the DAI in the Netherlands and subsequent superseding by ORCID/ISNI, [https://zenodo.org/record/7327505](https://zenodo.org/record/7327505)
- The gradual implementation of organisational identifiers (OrgIDs), [https://zenodo.org/record/7327535](https://zenodo.org/record/7327535)
- PIDs for research instruments and facilities: an emerging PID domain in need of coordination, [https://zenodo.org/record/7330372](https://zenodo.org/record/7330372)
- IGSN – building and expanding a community-driven PID system, [https://zenodo.org/record/7330498](https://zenodo.org/record/7330498)
- RePEc Author Service: An established community-driven PID, [https://zenodo.org/record/7330516](https://zenodo.org/record/7330516)
- Failed PIDs and unreliable PID implementations, [https://zenodo.org/record/7330527](https://zenodo.org/record/7330527)
- The role of research funders in the consolidation of the PID landscape, [https://zenodo.org/record/7258210](https://zenodo.org/record/7258210)
The team behind the study

○ Pablo de Castro
Physicist. Open Access Advocacy Librarian at the University of Strathclyde in Glasgow since Jan 2017. Technical Secretary of the Dutch non-profit association euroCRIS since Jan 2018. Former OpenAIRE project officer. Member of the EOSC Association Task Force for PID Policy and Implementation. Associate of scidecode science consulting.

○ Ulrich Herb
Sociologist & Information Scientist, since 2001 Open Access expert/ project manager/ head of the Publication and Research Support Department at Saarland University, board member of the learned society for Information Science in the German-Speaking countries. Associate of scidecode science consulting.

○ Laura Rothfritz
Research assistant and PhD candidate at the Berlin School of Library and Information Science at Humboldt University Berlin. Associate of scidecode science consulting.

○ Joachim Schöpfel
Professor for Information Science at the University of Lille and independent consultant.

The study

• **Analysis** of the current state of the PID landscape in the six Knowledge Exchange partner countries with a focus on the e-infrastructure for *the currently available PID entities (eg researchers, institutions, outputs, etc)* and *new PIDs (eg conferences, research equipment, facilities)*

• Data collection via **literature study & expert interviews**

• These fed into:
  
  • the production of the seven **case studies** highlighting issues of risk and trust in the PID infrastructure, and

  • the formulation of **recommendations** for good practice and on the best possible strategic and operational paths to achieve a well-functioning PID infrastructure
# The interviewees

<table>
<thead>
<tr>
<th>Name</th>
<th>Organization</th>
<th>Country</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathias Astell</td>
<td>Hindawi</td>
<td>GBR</td>
<td>PID Manager</td>
</tr>
<tr>
<td>David Aymonin</td>
<td>ABES</td>
<td>FRA</td>
<td>PID Authority</td>
</tr>
<tr>
<td>Geoffrey Bilder</td>
<td>CrossRef</td>
<td>GBR</td>
<td>PID Service Provider</td>
</tr>
<tr>
<td>Matt Buys</td>
<td>DataCite</td>
<td>GBR</td>
<td>PID Service Provider</td>
</tr>
<tr>
<td>Maria Cruz</td>
<td>NWO</td>
<td>NL</td>
<td>PID Manager</td>
</tr>
<tr>
<td>John Doove</td>
<td>SURF</td>
<td>NL</td>
<td>PID User</td>
</tr>
<tr>
<td>Nathalie Fargier</td>
<td>CNRS</td>
<td>FRA</td>
<td>PID Owner</td>
</tr>
<tr>
<td>Martin Fenner</td>
<td>formerly Technical Director at DataCite, involved in the FREYA project</td>
<td>GER</td>
<td>PID Manager</td>
</tr>
<tr>
<td>Stephanie Hageman-Wilholt</td>
<td>TIB Hannover/ConflIDent</td>
<td>GER</td>
<td>PID Authority</td>
</tr>
<tr>
<td>Juha Hakala</td>
<td>URN representative, National Library of Finland</td>
<td>FIN</td>
<td>PID Service Provider</td>
</tr>
<tr>
<td>Lars Holm Nielsen</td>
<td>Zenodo</td>
<td>CHE</td>
<td>PID Owner</td>
</tr>
<tr>
<td>Karen Hytteballe Ibanez</td>
<td>DTU - Technical University of Denmark</td>
<td>DNK</td>
<td>PID User</td>
</tr>
<tr>
<td>Jens Klump</td>
<td>IGSN</td>
<td>GER</td>
<td>PID Service Provider</td>
</tr>
<tr>
<td>Rachael Larmey</td>
<td>CrossRef</td>
<td>GBR</td>
<td>PID Service Provider</td>
</tr>
<tr>
<td>Dan Smith</td>
<td>Wellcome Trust</td>
<td>GBR</td>
<td>PID Owner</td>
</tr>
<tr>
<td>Mark van de Sanden</td>
<td>SURF, systems architect</td>
<td>NL</td>
<td>PID Authority</td>
</tr>
<tr>
<td>Herbert Van de Sompel</td>
<td>DANS</td>
<td>NL</td>
<td>PID User</td>
</tr>
<tr>
<td>Peter Verhaar</td>
<td>Leiden University</td>
<td>NL</td>
<td>PID Owner</td>
</tr>
</tbody>
</table>
Some (selective) findings

• PIDs are considered socio-technical infrastructures. Trust in organisations or individuals seems to be more important for the acceptance of PIDs than the technology used, as the risks associated with the technology are considered amorphous.

• Predominantly mentioned: well-established PIDs such as DOI, ORCID and ROR, to a lesser extent emerging PIDs (funder and grant IDs, RAiDs, ConfIDs), standards like URN and schemes like ARK.

• Main benefits: Interoperability, value-added services, availability/interconnectivity of rich metadata.

• Dichotomy of ‘technical’ (bottom-up, researcher driven) and ‘admin-oriented’ PIDs (top-down, uptake driven by institutions, publishers and research funders).

• Open source and open data are a key feature for trust and reliability.

• Establishing a community of PID users is a key factor for success and trustworthiness.

• The implementation of PIDs requires a strategic analysis.

• Significant PID landscape fragmentation and competing initiatives, though this is not necessarily seen as a major issue.
Rationale for Grant ID adoption by funders

- Risk of duplicate full grant numbers across funders

Grant finder

Europe PMC funder grants search results

Grant ID: 207467

Search again

1-2 of 2 results

Deep Earth Chemistry of the Core (207467)
Dr Badro, J, INSTITUT DE PHYSIQUE DU GLOBE DE PARIS, European Research Council 2008-2013 | Expired
8 publications | 6 free full text articles

Connecting causes and immune consequences of infection-induced metabolic change (207467)
Professor Dionne, Marc, Imperial College London, Wellcome Trust 2018-2023
7 publications | 6 free full text articles

1-2 of 2 results
Rationale for Grant ID adoption by funders

- Risk of duplicate full grant numbers across funders
- Typos frequently made by authors when acknowledging funding in their manuscripts
- ‘Harnessing the power of the PID Graph’: machine-readable grant IDs will allow easy interlinking across entities (esp btw grants and publications, datasets and other research results)
The (Currently) Fragmented PID Landscape

1. ‘Technical’ vs Admin-Oriented PIDs

- **Technical PIDs**: Researcher-led, bottom-up implementation, little if any direct involvement from research funders, research-performing organisations (RPOs) or national offices
  - **Examples**: PIDs for instruments and facilities, IGSNs, ISRCTNs (International Standard Randomised Controlled Trial Numbers), Biomedical IDs (GenBank etc)

- **Admin-Oriented PIDs**: led (generally in a top-down fashion) by national offices, RPOs, publishers and (some) funders, all of whom directly reap the rewards. Little researcher involvement*, even awareness (as a rule)
  - **Examples**: DOIs for publications/datasets, ORCID, RORs, Grant IDs, RAIDs, etc

* Researchers often see these as unwanted additional bureaucracy
Pavel Zbornik: „At present Wikidata ID is the most suitable ID for organisations in terms of coverage“

(interview by Pablo de Castro)

[with additional input by Stephane Ndong at the EC RTD]

What led the EC to take this step? What main use cases do you see for grant IDs?

There was an internal discussion on the use of grant IDs which resulted in an agreement that the EC should issue grant IDs for Horizon Europe. The overall purpose for the Commission is to contribute to the general development of PIDs in a mid- and long-term timescale. The idea underpinning this effort is that the availability of PIDs will become an important factor in the steering and evaluation of R&I Policies.

One major use case for grant IDs is linked to the evaluation and monitoring of the impact of the framework programme. The expectation is that better links between grants and related publications – plus other research outputs such as patents, prototypes, software components, etc – will simplify the analysis when the link is declared via a PID instead of as free-form text.
The (Currently) Fragmented PID Landscape

2. Competing Technical Solutions

¿CÓMO ADOPTAR ARK-CAICYT?

Etapa inicial: primeros adoptantes

- Disponer de un ISSN en línea provisto por la agencia argentina
- Publicaciones con URL únicas por recurso o parte
- Dos años de antigüedad en su versión en línea (etapa inicial de adopción)
- Adhesión a políticas de Acceso Abierto (Acceso inmediato, sin requerimientos de registro, suscripción o pago a los recursos)
- Nombrar un editor responsable para la gestión de su ARK-CAICYT
The (Currently) Fragmented PID Landscape

3. ‘Community’: an ambiguous concept

- PIDs and services associated with them need to be perceived as valuable and be in turn promoted by "the community"
Some recommendations for RPOs

1. **Make sure you are represented in** – or at least informed about – national-level coordination initiatives

2. **Consider** the possibility of drafting an **institutional PID policy**

3. **Raise awareness** of the existing and emerging PID landscape among institutional researchers, including prompting them to use the appropriate ones

4. **Be aware of your key role** in the implementation of specific, admin-oriented PIDs

5. **Include as many PIDs as possible** in your research information management systems such as institutional repositories and CRIS systems (plus any other institutional system that feeds these)

6. **Be aware of technical PIDs** directly emerging from researcher communities in a bottom-up fashion

7. **Stay informed about (still to come) mechanisms** to issue (and share and use) institutional PIDs such as RAIoDs or PIDINSTs
Thanks!

Questions?

Pablo de Castro

euroCRIS Secretary

Open Access Advocacy Librarian
University of Strathclyde
pablo.de-castro@strath.ac.uk
http://orcid.org/0000-0001-6300-1033
@pcastromartin