

# Australian National Persistent Identifier (PID) Strategy

Version 1 (July 2023)

## Preamble

Persistent Identifiers (PIDs) are a core component of a strong digital infrastructure ecosystem that accelerates research and innovation. PIDs can directly contribute to:

- Improving research efficiency, productivity and reduction in administrative burden
- Improving research reproducibility, provenance and attribution and hence research integrity
- Determining the value of, and return of investment in, research investments
- Understanding the relationship between the elements within the national research and innovation ecosystem with a view to optimising the system
- Tracking research engagement, translation and impact.

'PIDs are unique alpha-numeric codes that positively identify entities such as people, places, and things. In addition, they are connected to registries of information about those entities, known as metadata, that enable robust linking to and between those entities, thereby establishing provenance and attribution, as specified by the [FAIR principles](#). Further, PIDs contribute to research integrity by precisely identifying the resources used to conduct research and the outputs that result from it. The linking of PIDs to each other supports identification of scientific concepts across the system, driving research innovation and providing the evidence base for informed strategic decision-making at the institutional and national levels'.

The above quote is an excerpt from the 2022 report, commissioned by the Australian Research Data Commons (ARDC) and the Australian Access Federation (AAF), [Incentives to invest in identifiers: a cost-benefit analysis of persistent identifiers in Australian research systems](#). The report found significant potential savings to the Australian research system and that to realise these benefits will require a coordinated, comprehensive and collaborative approach to PIDs. The ARDC's approach is to convene a broad and open national discussion that will result in a widely understood national PID strategy and an accompanying roadmap.

This draft Strategy has been developed by the ARDC through a series of national workshops, together with strategic advice received via a *National PID Strategy Taskforce* comprised of senior stakeholders. It will be refined and accompanied by a five year Roadmap through an extensive, open, co-design consultation and engagement process with the sector in the second half of 2023.

The Strategy and Roadmap are not binding top-down policies. They are rather a stakeholder-driven vision of a future research system optimised through use of PIDs. They seek to provide a shared framework to inform co-investment and policy development.

We are delighted to share this draft Strategy with you and invite your feedback and engagement with the Strategy and Roadmap development process. To learn more, please visit:

<https://ardc.edu.au/project/australian-national-persistent-identifier-pid-strategy-and-roadmap/>

## Introduction

Australia's national research information infrastructure maximises the capability of Australia's research and innovation system to contribute to economic outcomes, national security, social wellbeing and environmental sustainability. Research is largely project-based and increasingly international and multidisciplinary. This is a necessity as we seek to address challenges such as climate change, food security and improving population health.

Persistent Identifiers (PIDs) are a core component of a world class, global digital information ecosystem. They enable us to manage the fragmented, complex, distributed and federated research and innovation ecosystem by providing a universal, machine-readable, interoperable method to uniquely identify and connect entities such as researchers and innovators, funders, organisations, articles, datasets, software, instruments and samples. These connections allow us to derive insights which will accelerate Australia's research and innovation ecosystem.

This Strategy is the culmination of a coordinated, comprehensive and collaborative process led by the Australian Research Data Commons (ARDC). It has brought together a National PID Strategy Taskforce and key stakeholders across the national research, innovation and impact ecosystem. The Strategy is underpinned by a National PID Roadmap.

The ultimate beneficiary of the Strategy and Roadmap will be the Australian people. By strengthening our digital information ecosystem, we strengthen the quality of our research and our capacity for effective research engagement, innovation and impact. In recognition that research is increasingly international, the Strategy and Roadmap will maximise national outcomes, adopt international best practices and contribute to the development of international persistent identifiers and standards.

## The Challenges

Managing the ever increasing amount and complexity of data about and stemming from research is essential to research and innovation, and to our ability to understand and improve the ecosystem. Referring to the data produced through research, the [National Research Infrastructure Roadmap 2021](#) found that *Exponential growth in data across all disciplines will be a critical challenge for NRI over coming years, highlighting the need for integration of computing and data infrastructure and the maintenance of a strong digital infrastructure ecosystem.*

Deriving insights from this data will increasingly require use of compute tools, requiring that data is FAIR (Findable, Accessible, Interoperable and Reusable) and is consistent with the CARE Principles for Indigenous Data Sovereignty (Collective benefit, Authority to control, Responsibility, Ethics).

PIDs are core to maximising data use and impact. Currently, a large portion of the data elements which are integral to research are siloed, difficult to find or access, often leading to replication and duplication of data collection

efforts and, even more importantly, reducing the opportunity for cross-disciplinary research to address the complex challenges we face. For example, if we are to address the impact of climate change on our nation's environment we need the ability to share the observations, data and insights gained through the many environmental impact assessments undertaken across multiple levels of government and by research agencies. Whilst PIDs cannot make people share data, they can make it easy for them to do so and in a way that can create real value, improve research integrity and attribution. Imagine if we could share the data collected through environmental impact studies through appropriate use of PIDs, accelerating our research capability and ultimately its impact? Imagine the savings we could make through reducing duplication of effort?

This lost opportunity, and waste of effort, is also found in the management of data about research, where data cannot be used to effectively understand the relationships across the research and innovation ecosystem to optimise the system, and where the same data is collected many times over. For example, we cannot easily understand which research investments yield which outcomes, which research grants have the most impact, or how well equipped we are to undertake the research most needed to meet national priorities. Each year the Federal government invests billions of dollars into research and innovation across many different programs and government agencies. Imagine if we could trace the impact of these grants into research and innovation outcomes through use of linked persistent identifiers? Imagine if researchers and research agencies could easily identify what grants are available, across all granting agencies to research pressing national priorities? Quantifying this problem, a 2022 report commissioned by the ARDC and AAF, [Incentives to invest in identifiers: a cost-benefit analysis of persistent identifiers in Australian research systems](#), found that up to 38,000 person days of effort per year could be saved through use of persistent identifiers (PIDs).

## The Opportunity

Persistent identifiers can maximise the value of the investments we are making in our research and innovation ecosystem by:

1. Improving research quality through reducing duplication, improving research reproducibility, provenance and attribution;
2. Catalysing innovation and impact through use of PIDs which apply across sectors, disciplines and borders;
3. Increasing the efficiency of the research and innovation ecosystem, including a reduction in administrative burden; and
4. Optimising the research and innovation ecosystem through better understanding the relationship between the elements of the ecosystem, including between investment, use, outputs, outcomes and impact.

Adoption of ORCID in Australia has already demonstrated the value of PIDs. The Australian Research Council's use of ORCIDs to populate research grant applications was estimated by the *Incentives to invest in Identifiers* analysis report to have saved researchers up to three or four days per grant application. These benefits have been realised as a result of a national approach to solving the common problem of needing to correctly identify individual researchers and accurately link them to their works: there was a broad, open and inclusive consultation process that included collaborative development and launch of the model for the Australian ORCID Consortium.

There is growing momentum nationally and internationally to better leverage data stemming from research and about research and the potential offered by PIDs. Internationally, many nations, including the United Kingdom and Canada, are well advanced in development of a national approach to PIDs to underpin their research and innovation ecosystems.

## The Vision

Accelerate Australian research quality, efficiency and impact through universal use of connected persistent identifiers.

## The Strategy

The vision will be realised through a national, coordinated, and collaborative approach across the research and innovation ecosystem. The Strategy will enable implementation of PIDs within research organisations, across all levels of government and in business and industry (where they are partners and contributors to research and innovation).

### Improve research quality and efficiency

- 1. Increase the Findability, Accessibility, Interoperability and Reuse of inputs to research*  
to research include potential research partners, related research projects, grants, facilities, data, samples and observations. By increasing a researcher's ability to discover relevant inputs, we will reduce duplication, increase research productivity and potentially create impactful partnerships.
- 2. Increase the Findability, Accessibility, Interoperability and Reuse of research outputs*  
The outputs from research include data, reports, methods, non-traditional research outputs and publications. By increasing the discoverability of these outputs, we can improve future research, create new kinds of research, reduce duplication and improve the timeliness of translating research into impact.
- 3. Improve research reproducibility, provenance and attribution whilst minimising administrative burden, enabling researchers to spend more time on research*

By linking researchers to projects, to grants, to organisations, to equipment, we can improve research reproducibility, provenance and attribution. Through the use of PIDs researchers can minimise the amount of time they spend on reproducing the same data many times over, releasing valuable time back into research.

### Optimise the national research and innovation ecosystem

- 4. Improve our ability to evaluate research quality, impact and evidence of public benefit*  
By linking elements across the research and innovation ecosystem through PIDs, we can more efficiently assess the performance of our research institutions with respect to research quality thresholds, better trace the impact of the research and through this, more readily explain the public benefit of the research.
- 5. Optimise our ability to understand the impact of research inputs such as grants and investment in research infrastructure*

By linking elements across the research and innovation ecosystem through PIDs we can better understand the relationship between the elements of the ecosystem, including between investment, use, outputs, outcomes and impact.

6. *Improve our ability to map Australia's research capability*

Australia needs the research capability to address our national priorities. PIDs for people, organisations, research services and infrastructure would enable us to map our research capability against national priorities and invest appropriately to address gaps.

## Implementation

Implementation of the Strategy will be in accord with the accompanying five-year Roadmap which will be refreshed and extended annually. The Roadmap outlines the criteria which inform priority setting, the actions to be taken, by who and by when, to achieve the intent of the Strategy. Oversight of the Roadmap will be through a cross-sector Advisory Taskforce.