ARDC

Australian Research Data Commons

Australian National Persistent Identifier (PID) Strategy 2024

February 2024

Acknowledgement of Country

We acknowledge the traditional custodians throughout Australia and their continuing connection to, and deep knowledge of, the land and waters. We pay our respects to Elders both past and present.

Acknowledgements

Publication Information

The Australian National Persistent Identifier (PID) Strategy and Roadmap will benefit the Australian people by strengthening our digital information ecosystem, the quality of our research and our capacity for effective research engagement, innovation and impact.

Australian National Persistent Identifier (PID) Strategy 2024

Persistent Identifiers (PIDs) are a core infrastructure component of a world-class, global digital information ecosystem. They provide a universal, machine-readable, interoperable method to uniquely identify and connect entities such as researchers and innovators, funders, organisations, articles, datasets, projects, software, instruments and samples. This allows us to manage the fragmented, complex, distributed and federated research and innovation ecosystem, and derive insights that will accelerate it for Australia.

The Australian National Persistent Identifier Strategy is the culmination of a coordinated, comprehensive and collaborative process led by the Australian Research Data Commons (ARDC). It has been informed by extensive consultation through national workshops, working groups, webinars and via an open call for individual and group submissions. A National PID Strategy Taskforce comprised of senior stakeholders provided strategic advice and guidance. Released as a draft in July 2023, the Strategy has been refined through extensive and open co-design, consultation and engagement with the sector.

The Strategy is accompanied by a Roadmap, which provides an actionable national collaborative plan and a set of tools to support implementation of the Strategy. Together, they will benefit 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 through adopting international best practices and contributing to the development of international persistent identifiers and standards.

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 that helps relevant organisations with co-investment and policy development.

We are delighted to share this Strategy with you, and we invite you to engage with the accompanying Roadmap. To learn more, please visit - https://bit.ly/pids-strategy

Introduction

Australia's national research information infrastructure maximises our research and innovation system's capability to contribute to the economy, national security, social wellbeing and environmental sustainability. Research is largely project-based and increasingly international and multidisciplinary. It is necessarily so as it addresses challenges such as climate change, food security and population health.

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

- research efficiency, productivity and a reduction in administrative burden

- research reproducibility, provenance, attribution and, in turn, integrity

- determining the value of, and return of investment in, research investments

- understanding how elements of the national research and innovation ecosystem relate to one another with a view to optimising the system

- tracking research engagement, translation and impact.

"

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.

Brown, J., Jones, P., Meadows, A., & Murphy, F. (2022). Incentives to invest in identifiers: A cost-benefit analysis of persistent identifiers in Australian research systems. *Zenodo. https://doi.org/10.5281/zenodo.7100578*

The 2022 report *Incentives to invest in identifiers: a cost-benefit analysis of persistent identifiers in Australian research systems* (hereafter *Incentives*), commissioned by the ARDC and the Australian Access Federation (AAF), notes:

> 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 report finds significant savings to the Australian research system if a coordinated, comprehensive and collaborative approach to PIDs is in place — one that brings together the key stakeholders across the national research, innovation and impact ecosystem.

The Challenges

Managing the ever-increasing amount and complexity of data about research and data 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 2021 National Research Infrastructure (NRI) Roadmap finds that "[e]xponential 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 depend on compute tools, demand that data is FAIR (Findable, Accessible, Interoperable and Reusable), and require that it is consistent with the CARE Principles for Indigenous Data Sovereignty (Collective benefit, Authority to control, Responsibility, and Ethics).

PIDs are core to maximising data use and impact. Currently, a large portion of the data elements integral to research are siloed and difficult to find or access. This often leads to replication and duplication of data collection efforts and, even more critically, hampering the cross-disciplinary research needed 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 to be able to share the observations, data and insights gained through the many environmental impact assessments undertaken across multiple levels of government and by research agencies. While PIDs do not compel one to share data, they can make it easy to do so and in a way that creates real value, improves attribution and promotes research integrity. Through appropriate use of PIDs, we will have greater research capability, our work will have more impact, and savings can be made with less duplication of effort.

This lost opportunity, and waste of effort, is also found in the management of data about research. Data often cannot be used to effectively understand how elements of the research and innovation ecosystem relate to one another in order to optimise the system, and the same data is often collected many times over.

For example, we cannot easily understand what outcomes a particular research investment has had, which research grants have the most impact, or how prepared we are to undertake the research most needed to meet national priorities. Each year, the federal government invests billions of dollars in research and innovation through various programs and government agencies. Through use of linked PIDs, we could trace the impact of these grants on research and innovation outcomes, and researchers and research agencies could easily identify what grants are available from the array of granting agencies for research addressing national priorities. Quantifying this problem, *Incentives* finds that up to 38,000 person days of effort could be saved per year through use of PIDs.

While PIDs do not compel one to share data, they can make it easy to do so and in a way that creates real value, improves attribution and promotes research integrity

Images — Nick Starichenko - 379768765 / AdobeStock.com

The Opportunity

PIDs can maximise the value
of the investments in our research
and innovation ecosystem by:

1. enhancing research through reducing duplication and improving research reproducibility, provenance and attribution

2. catalysing innovation and impact through use of PIDs, which apply across sectors, disciplines and borders

3. making the research and innovation ecosystem more efficient, including through reducing administrative burden

4. optimising the research and innovation ecosystem through facilitating understanding of how elements of the ecosystem, including investment, use, outputs, outcomes and impact, relate to one another.

Adoption of ORCID in Australia has already demonstrated the value of PIDs. *Incentives* estimates that by using ORCIDs to populate research grant applications, the Australian Research Council (ARC) has saved researchers up to 3 or 4 days per grant application. The problem of needing to correctly identify individual researchers and accurately link them to their works had been common; the benefits of ORCIDs were realised thanks to our national approach to the problem – a broad, open and inclusive consultation process that included the 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 to harness the potential offered by PIDs. Many countries, including the UK and Canada, are well advanced in developing a national approach to PIDs to underpin their research and innovation ecosystems.

Image — Renee Nowytarger - AusTraits / ARDC

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, comprehensive and collaborative approach across the research and innovation ecosystem. PIDs will be implemented in 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**

   Inputs 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 create opportunities for impactful partnerships.

2. **Increase the Findability, Accessibility, Interoperability and Reuse of research outputs.**

   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 while minimising administrative burden, enabling researchers to spend more time on research**

   By linking researchers to projects, grants, organisations and 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

1. **Improve our ability to understand the impact of research inputs and evaluate research quality, impact and evidence of public benefit.**

   By linking elements across the research and innovation ecosystem through PIDs, we can better understand the relationship between the elements of the ecosystem, including investment, use, outputs, outcomes and impact. We can more efficiently assess our research institutions' performance 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.

2. **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 enabled by an accompanying Roadmap, which will be regularly refreshed and extended. The Roadmap is an actionable national collaborative plan and a set of tools for achieving the Strategy's intent. It outlines the criteria for setting priorities, the actions to be taken, by who and by when. Stakeholders in the research and innovation ecosystem will accelerate the quality and impact of Australian research by collaboratively contributing to a national PID system. The Roadmap is framed by a cycle of continuous improvement.

ARDC

Australian Research Data Commons

At the ARDC, we drive development of world-class national digital research infrastructure that gives Australian researchers competitive advantage through data and supports research impact.

The ARDC is Australia's leading facility for research data infrastructure. We facilitate access to research data sets and tools from academia, industry and government for all Australian researchers.

We run programs and form partnerships that ensure Australian researchers are internationally competitive through having access to high-quality data assets, platforms, infrastructure, policies, people and training to transform our lives.

Solving society's greatest challenges takes the collective efforts of society. Through our collaborations and partnerships — national and international — we are ensuring that valuable data and software assets are developed, made accessible and sustained for everyone.

The ARDC is enabled by the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS).

FOLLOW

- @ardc_au
- australian-research-data-commons
- ARDC_AU
- Subscribe to the ARDC Connect Newsletter
  https://ardc.edu.au/subscribe/

CONTACT

- ardc.edu.au
- +61 3 9902 0585
- contact@ardc.edu.au