

# MORE+BRAINS

The case for investment in a UK  
persistent identifier strategy:

Resilience, insight, and  
leadership in global research  
and innovation

Prepared by: Josh Brown of MoreBrains Cooperative

Date: 29/11/2021

Version 2.

## Executive summary

Modern research operates in an era of international collaboration. At the same time, competition between research powers for talent and impact is growing. In this context, shared global research information systems are vital to support cooperation and to enable an accurate strategic analysis of the research landscape.

Persistent Identifiers (PIDs) are unique digital codes that uniquely identify the entities (such as people, project, funding grants, and outputs) that make up much of the research landscape. They are often international in reach, and operate as bridges between information silos. A set of high-priority PIDs has been the focus of a multi-year programme in the UK, sponsored by Research England and commissioned by Jisc<sup>1</sup>.

That programme has shown that there is both a need and a strong business case for investment in a PID support framework for the UK research and innovation sector. Without comprehensive adoption and use of PIDs, organisations large and small will be disadvantaged in the modern, digital-first, research environment.

This document sums up the findings of this programme, and calls for a national investment in activities to implement PIDs in alignment with the UK's strategic needs. The steps necessary are:

1. Create a consortial network to lower barriers to and costs of PID adoption
2. Community education and promotion to drive adoption of PIDs to leverage network effects
3. Understand critical information pathways between funders, institutions, and content publishers
4. Establish technical and social requirements for systematic exchange and reuse of information across stakeholder groups

Extensive research and consultation has shown that three components are necessary to deliver these needs. Two of these, national access to and support for key PIDs, and a national coordination body to guide and oversee work in this space, are in place or under development. The third component is a dedicated team of technical, educational, and communications specialists hosted at a trusted UK institution to drive forward the PID agenda<sup>2</sup>. We are calling for funding to support the creation of this team, with a long enough span to enable the community to develop a model for supporting the work on an ongoing basis.

This document sets out some of the central arguments for why such an investment is needed now. These include:

- PIDs are required to deliver a range of policy priorities and can help to align and join up work across the drive to deliver them (section 2)
- PIDs have the power to save time and money and boost the return on investments in research (section 3)

---

<sup>1</sup> See Appendix A: Reports and details of work completed

<sup>2</sup> See Appendix B: Theory of change

- PIDs provide a means for the UK to maintain its competitive advantage in the hyper-connected global research landscape (section 4)
- The power of information systems has been vital in bringing the UK through a number of recent crises. The resilience of the UK's knowledge economy is clearly dependent on the resilience of these systems, and PIDs provide a means of strengthening these systems against future shocks and the risk of information loss (section 5)
- In an age of global research collaboration, open, sustainable, international information systems are vital both for the trustworthiness of our research network and for our ability to analyse and adapt strategically (section 6)

## 1. Introduction

Modern research and innovation are increasingly global, digital-first, and hyperconnected. Having passed through three ages, the individual, the institutional, and the national, research is now in its fourth: the age of international collaboration.<sup>3</sup> In this age, research relationships and interactions are ferociously complex, and research production has reached a vast scale.

Complexity at such a scale creates new and daunting challenges for anyone seeking to navigate the modern research landscape. To understand and evaluate information we must understand its provenance, origins, and relationships, but this essential information is often fragmentary, spread across multiple potential sources, and hard to verify. In this context, persistent identifiers (PIDs) have emerged as a critical tool for mapping the world of research and innovation (R&I)<sup>4</sup>.

A PID is a unique digital alphanumeric code which serves as a long-lasting set of coordinates for a single entity (such as a person, a journal article, a dataset, or a gene sequence). PIDs provide both a link to the entity, and information (metadata) about it. The relationships and links between PIDs can enable new navigation and analysis for R&I.<sup>5</sup>

The UK has been leading the world in developing a roadmap for comprehensive PID adoption<sup>6</sup>. Jisc<sup>7</sup>, sponsored by Research England<sup>8</sup>, have led a series of projects with the goal of providing a consistent information infrastructure to enhance the resilience, efficiency, and excellence of the UK research and innovation sector.

Collectively, these projects found that improving the interconnectedness and reusability of research-related metadata with PIDs will increase the accuracy and completeness of research information, reduce administrative burden for researchers and organisations alike, and enable better decision-making and prioritisation of investments in R&I.

<sup>3</sup> Adams, J. The fourth age of research. *Nature* 497, 557–560 (2013). <https://doi.org/10.1038/497557a>

<sup>4</sup> See Appendix C: Case Study - PIDs for projects

<sup>5</sup> Klump, J., Murphy, F., Weigel, T. and Parsons, M.A., 2017. Editorial: 20 Years of Persistent Identifiers – Applications and Future Directions. *Data Science Journal*, 16, p.52. DOI: <http://doi.org/10.5334/dsj-2017-052>

<sup>6</sup> Brown, Josh (2020) Developing a persistent identifier roadmap for open access to UK research. <https://repository.jisc.ac.uk/7840/>

<sup>7</sup> <https://www.jisc.ac.uk/>

<sup>8</sup> <https://www.ukri.org/councils/research-england/>

## 2. Prioritising PIDs for the UK

The value and potential of PIDs is already widely recognised. The Universities UK Open Access Coordination Group, chaired by Professor Adam Tickell, recommended the use, and in some cases mandation, of PIDs.<sup>9</sup> UKRI's open access policy calls for the extensive use of PIDs for journals, articles, authors, grants, funders, organisations, and projects.<sup>10</sup> The FAIR principles (for the Findability, Accessibility, Interoperability, and Reusability of research data) rely on PIDs to achieve their aims.<sup>11</sup> The UK's Department for Business, Energy, and Industrial Strategy released a policy paper on reducing bureaucratic burden in research, innovation and higher education in 2020, which called for "[s]topping multiple asks for data or information that already exists elsewhere e.g. in ORCID, CrossRef, DataCite and Companies House."<sup>12</sup>

All these policy documents reflect a simple truth: PIDs are a critical dependency in a modern, digital-first, research system.

Community consultation found that the top five priorities are PIDs for grants, people, projects, organisations, and outputs. Existing open, non-proprietary, community-governed, international PID systems can be used for each of these entities, and we have laid out a roadmap for their implementation.<sup>13</sup> It is also clear from our research that the use of these PIDs is needed throughout the research lifecycle, from funding applications to institutional research management to the publication, sharing, and reuse of outputs.<sup>14</sup> When so many policy priorities rely on PIDs, a coherent, well-resourced national strategy to align goals and deliver PIDs is in order.

## 3. Missing PIDs, missing efficiencies

A recent cost-benefit analysis<sup>15</sup> evaluated the case for increased investment in and community support for PID integrations. It showed that PIDs save money by increasing metadata re-use, automating processes, and improving the aggregation and analysis of information about research. Where those savings are being realised, they are significant: for example, the UK's ORCID consortium has saved between £2.4M and £9.6M over its six-year lifetime; and Crossref's automated sharing of article metadata via ORCID records has saved the UK

---

<sup>9</sup> Tickell, A. (2018) Open Access to Research Publications: Independent advice. Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/774956/Open-access-to-research-publications-2018.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/774956/Open-access-to-research-publications-2018.pdf)

<sup>10</sup> UKRI Open Access Policy (2021). See especially Annex 2: Technical requirements for research articles. <https://www.ukri.org/wp-content/uploads/2021/08/UKRI-180821-UKRIOpenAccessPolicy-2.pdf>

<sup>11</sup> <https://www.force11.org/group/fairgroup/fairprinciples>

<sup>12</sup> Department for Business, Energy, and Industrial Strategy/ Department for Education. (2020) Reducing bureaucratic burden in research, innovation and higher education. Available at: <https://www.gov.uk/government/publications/reducing-bureaucratic-burdens-higher-education/reducing-bureaucratic-burdens-on-research-innovation-and-higher-education>

<sup>13</sup> Brown, Josh (2020) Developing a persistent identifier roadmap for open access to UK research. Available at: <https://repository.jisc.ac.uk/7840/>

<sup>14</sup> Brown, Josh and Meadows, Alice (2020) Persistent identifiers adoption and awareness survey report. <https://repository.jisc.ac.uk/8107/>

<sup>15</sup> Brown, Josh, Jones, Phill, Meadows, Alice, Murphy, Fiona, & Clayton, Paul. (2021). UK PID Consortium: Cost-Benefit Analysis (1.0). Zenodo. <https://doi.org/10.5281/zenodo.4772627>

between £1.26M and £5M since 2015. These findings point to the potential scale of savings to be gained from the comprehensive use of PID registries and services.

Conversely, where these efficiencies are not in place, time and money are being wasted, preventing the UK research, innovation, and higher education sector from maximising its impact. For instance, our work showed that a 2% increase in efficiency in academic R&D could bring £420M in benefits to the wider UK economy. A 2% efficiency gain is a conservative estimate of what could be achieved as a return on a relatively modest investment in PIDs for UK research and innovation.

## 4. Global challenges

The UK has been an international research heavyweight for centuries, but, as governments around the world prioritise the knowledge economy, a number of rising research powerhouses are now disrupting the established research order. For example, in 2014 China overtook the UK in global share of highly cited articles.<sup>16</sup> This success remains central to Chinese government plans, as is borne out by their Action Plan for the Excellence of Chinese STM Journals.<sup>17</sup> Countries from Brazil to Israel to China are all ramping up their research funding steadily.<sup>18</sup>

If we are to remain in the top tier of research nations whilst living within our means, the efficiency and connectedness of the UK research and innovation sector is critical. Rising research powers do not have the same workforce constraints that the UK does, and, unlike us, they can overcome inefficient systems by force of staff numbers. Instead, we need automation and joined-up systems to do this work for us.

Research by HM Treasury shows that ‘the value of the leading companies in the world today is not in their physical assets but in their IP, R&D, know-how and data – intangible or ‘knowledge assets.’<sup>19</sup> Intangible assets have the following characteristic features: spillovers, synergies, sunk costs, scaling up. PIDs provide a low-cost, high reward route to expediting higher innovation levels within the public sector.

In this light, the ‘fourth age of research’ is both a valuable opportunity to extend the reach and impact of UK research and a significant threat to the UK’s standing as a research powerhouse. Without better data about activities and results to support improved analysis and planning, we will not be equipped to steer the UK R&I sector into the future. For universities, knowing where your staff are going, and who your own and rival institutions are collaborating with is

---

<sup>16</sup> Elsevier. (2017) ‘International Comparative Performance of the UK Research Base – 2016’, Department for Business, Energy & Industrial Strategy (BEIS), UK.

<https://www.elsevier.com/research-intelligence/research-initiatives/BEIS2016>

<sup>17</sup> Impact Science (2021) ‘Navigating China’s Academic and Research Landscape: A guide for academic publishers and societies.’ Available at:

<https://www.impact.science/whitepaper/navigating-china-academic-research-landscape.html>

<sup>18</sup> Source: World Bank Data (2021)

<https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS?locations=BR-CN-IL>

<sup>19</sup>

<https://knowledgeasset.blog.gov.uk/2020/11/30/the-knowledge-economy-innovation-in-the-uk-public-sector/>

crucial. It enables them to manage the risks and benefits of international collaborations, as well as delivering the UK Government's R&D People and Culture Strategy goal of "[r]enewing the UK's position as a global leader in R&D in attracting, retaining and developing talented people, making sure R&D careers in the UK are appealing to talented individuals and teams both domestically and internationally".<sup>20</sup>

Other nations have now spotted the potential of PIDs and are seeking to leverage the work begun in the UK to develop their own PID strategies.<sup>21</sup> The 'first mover' advantage was hard won for the UK and can be easily lost.

## 5. Resilience now and in the future

As the global post-COVID recovery begins to come into view, it is more important than ever for the UK to be in shape for the next shock. Pandemic-related disruptions from the loss and redistribution of workforce to the dramatic shift to remote working have highlighted the importance of remote access technologies and digital resources, but they have also scattered research activities and outputs across networks in ways that make tracking them, understanding them, and preserving them extremely challenging. Large number of retractions of COVID-relevant papers came about because it was impossible to validate the data underpinning their findings.<sup>22</sup>

The rapid expansion of digital content is exposing the limits of our current information systems. Information is lost at each stage of production and transmission, and must then be replaced by commercial players who typically do not expose their processes, algorithms, or sources. These layers of missing information then hide the context needed for critical evaluation and undermine the ability of researchers and others to exercise their information literacy.

At the same time, researchers, research managers, and administrators have been pushed to breaking point by the demands of shifting a complex research bureaucracy and delivering the 2021 REF.

PIDs are not the only solution to these institutional crises, but they are an essential component in our toolkit for underpinning and future-proofing the UK R&I digital infrastructure.

## 6. Open and International

Top quality research and innovation are not produced within a national vacuum, see, for example climate research as well as our cultural heritage collections (the latter are also relevant to a number of Commonwealth and other countries). Although our priority is to

---

<sup>20</sup> Department for Business, Energy & Industrial Strategy (2021) R&D People and Culture Strategy. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1004685/r\\_d-people-culture-strategy.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1004685/r_d-people-culture-strategy.pdf)

<sup>21</sup> See for example: Cruz, M. & Tatum, C. (2021). NWO Persistent Identifier Strategy. Zenodo. <https://doi.org/10.5281/zenodo.4674513>

<sup>22</sup> See for example: Ledford, H. & Van Noorden, R. (2020) High-profile coronavirus retractions raise concerns about data oversight, Nature 582, 160 doi: <https://doi.org/10.1038/d41586-020-01695-w>

strengthen the UK's research sector, this can only be achieved through strategic participation in international initiatives such as the global PID systems. Research is itself global, and international systems are needed if the UK is to maintain visibility of its mobile, collaborative research workforce and its outputs and impacts.

Open PIDs, including ORCID<sup>23</sup> and ROR<sup>24</sup>, provide a venue for building trusted systems through responsible governance and promoting interoperability. They enable services to be developed that support community needs, provide accountability, and overall will speed up dissemination of critical outputs as provenance and credit frameworks mature. They also reduce the risk of a single point of failure causing irretrievable losses of information<sup>25</sup>. The community governance of these organisations provides a means for the UK to ensure that the services and tools they provide are fit for purpose and meet the UK's needs, as well as providing the transparency needed to assess their robustness and compliance with principles of sustainability and best practice.<sup>26</sup>

## 7. It is time for strategic investment in PIDs

For all contributors to and components of UK research and innovation to have PIDs, for those PIDs to have good quality, reliable metadata associated with them, and for all the players in the system to be equipped to take advantage of that metadata will take effort and investment at every level of the research system. Our research has shown that current levels of integration, adoption, and coverage of PIDs, are not sufficient to deliver the benefits outlined here and in our cost-benefit analysis.

If UKRI's open access policy is to be effective and capable of being implemented and monitored efficiently, and for the reduction of repetitive asks for information outlined in the government's policy paper to be achieved, PIDs need to be consistently and reliably present and accompanied by full metadata and a rich ecosystem of services.

At present there are notable gaps to be addressed. Smaller institutions face a disproportionate burden in implementing technical systems that leverage PIDs compared to the size of their research base. Some disciplines, especially in the Humanities, Arts, and Social Sciences, have felt deprioritised by PID providers, and have been underserved more generally by digital scholarly systems.

Fragmentary and inequitable access to PID resources is not an effective or fair way to deliver a future-ready, more efficient and resilient research and innovation base for the UK. Unless everyone in the system can benefit, then no one in the system will get the full potential benefits of PIDs. Every player in the research system, from funders to publishers and from researchers

---

<sup>23</sup> <https://orcid.org/>

<sup>24</sup> <https://ror.org/>

<sup>25</sup> Dappert, A., Farquhar, A., Kotarski, R. and Hewlett, K., 2017. Connecting the Persistent Identifier Ecosystem: Building the Technical and Human Infrastructure for Open Research. *Data Science Journal*, 16, p.28. DOI: <http://doi.org/10.5334/dsj-2017-028>

<sup>26</sup> See for example the Principles of Open Scholarly Infrastructure. <https://openscholarlyinfrastructure.org/>



to policy makers, has both a role to play in executing a successful PID strategy and a stake in its outcome.

Existing PID consortia, such as the DataCite and ORCID consortia led by the British Library and Jisc respectively, provide a basis for building technical capacity, pooling knowledge, and levelling the field to enable widespread access across domains and relative institutional resources. The Research Identifier National Coordinating Committee (RINCC) stands ready to provide strategic oversight and international coordination for the future of PID activities in the UK. UKRI, Jisc, and the many stakeholders who have participated in the programme, whether in research, task forces or the RINCC, have all invested significant amounts in bringing the roadmap to its current stage.

The return on all these investments will decay without a timely, targeted follow-up investment to unlock and deliver the benefits outlined in this strategic case and in the cost-benefit analysis. Work done so far has identified a costed set of interventions which should take the UK PID agenda to a sustainable, self-supporting level, capable of delivering the wider goals of a range of policy priorities for the UK.

Our research and community consultations make it clear that this is not going to come without strategic leadership and vision from the very top. This vision cannot be unlocked without a concerted, sector-wide commitment to the strength, health, and effectiveness of the UK research and innovation enterprise.

## Appendix A: Reports and details of work completed

Current and past activities set out in the theory of change diagram have yielded a body of knowledge to guide future investments and support activities.

1A: Brown, Josh (2020) Developing a persistent identifier roadmap for open access to UK research. <https://repository.jisc.ac.uk/7840/>

- Summarised community consultation that led to the list of five priority PIDs
- Emphasises need for open and community-governed solutions
- Set out necessary steps to the creation of a robust national PID strategy

1A: Brown, Josh, & Meadows, Alice. (2021). UK "multi-PID consortium" business case. Zenodo. <https://doi.org/10.5281/zenodo.4760886>

- Laid out the value PIDs could offer to funders/policy makers, institutions, researchers, and to commercial 'academia-adjacent' players
- Highlighted the steps needed to create a strong support framework to drive PID adoption and coverage to levels that would deliver value

1B: Brown, Josh and Meadows, Alice (2020) Persistent identifiers adoption and awareness survey report. <https://repository.jisc.ac.uk/8107/>



- Showed that PIDs are sought by the community to improve system interoperability and to generate new insights into research practices and outcomes
- Revealed that technical barriers and costs of PID integrations are seen as too high at present, slowing PID adoption
- Pinpointed critical workflows in the R&I enterprise which can be optimised using PIDs
- Set out the need for a clear value proposition and senior buy-in to drive PID adoption and coverage upwards

1B and 1C: Brown, Josh, Jones, Phill, Meadows, Alice, Murphy, Fiona, & Clayton, Paul. (2021). UK PID Consortium: Cost-Benefit Analysis (1.0). Zenodo.

<https://doi.org/10.5281/zenodo.4772627>

- Calculated the time and financial losses being incurred by inefficiencies in the research system
- Highlighted the need for strong, high-level leadership to address systemic challenges equitably

2A: Murphy, Fiona and Jones, Phill (2020) Jisc PID Agency: Focus Group Report.

<https://repository.jisc.ac.uk/8166/>

- Established broad agreement on the PID strategy principles across UK and PID providers
- Highlighted the range of PID adoption across the sector, indicating where the most and least progress has been made
- Outlined the need for clear outreach, strong governance models, and sequential interventions directed at a range of stakeholders

2B: Establishment of Research Identifier National Coordinating Committee (RINCC)

<https://rincc.org.uk/>

- Combined membership across the majority of open research-supporting organisations in the UK
- Provide strategic oversight, troubleshoot barriers, provide national and international leadership
- Participate in evolving business models, forming consortia, supporting PID strategy initiatives

2B and 2C: Development of RINCC terms of reference

- <https://docs.google.com/document/d/1-TYRAFBIF5207ugs1xJfc7saasU7o-dfi66qjZ8rmZA/edit?usp=sharing>
- Meet twice yearly, self-seeding membership, decide sub-projects and areas of concern

2C: Research Data Alliance national PID strategies working group.

<https://www.rd-alliance.org/groups/national-pid-strategies-wg>

- Newly formed working group, co-Chaired by UK and Australia
- Mapping commonalities across national agencies, generating best practices and common standards
- Identifying governance and consortia models, leveraging investment opportunities

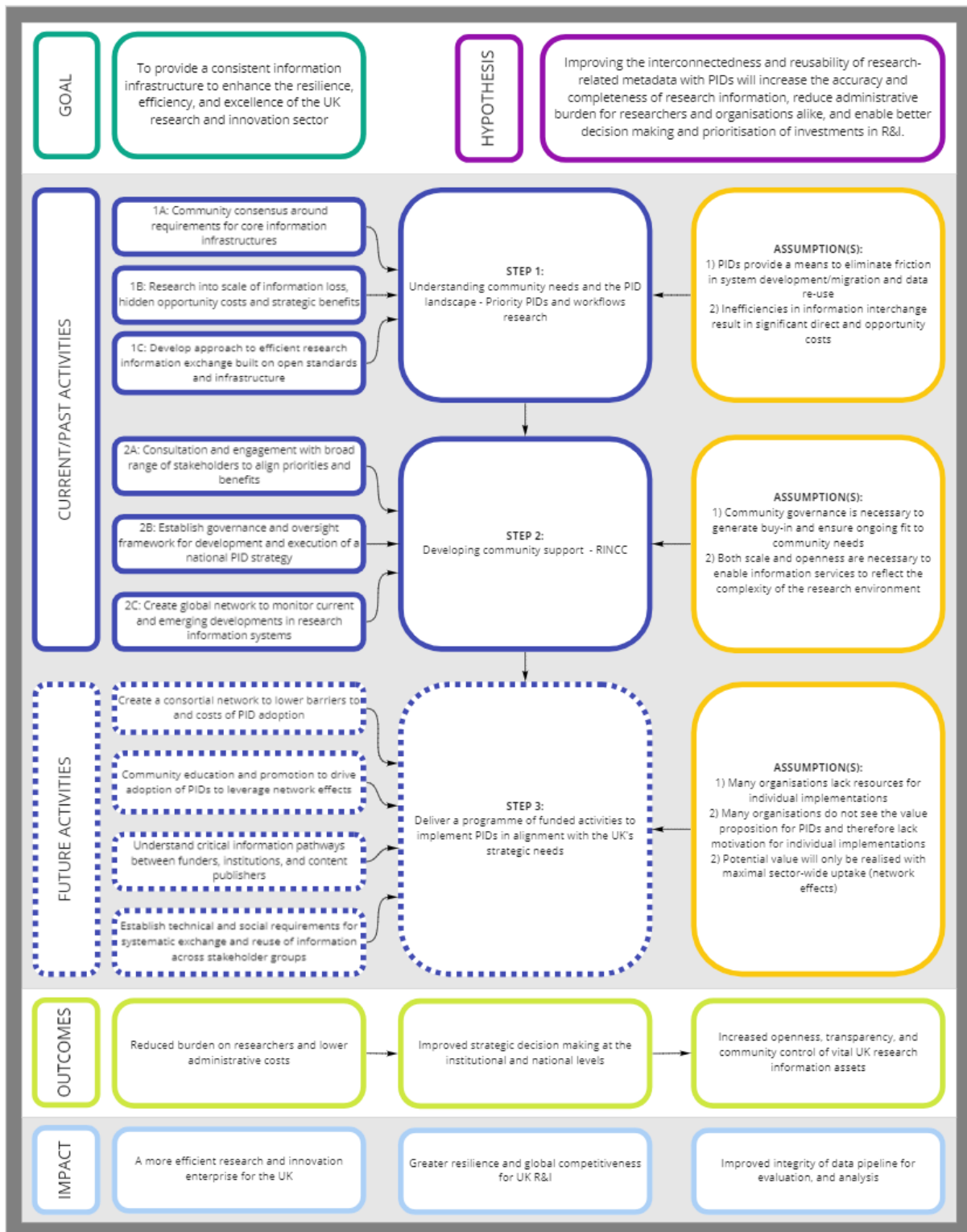
## Appendix B: Theory of change

The hypothesis underpinning an investment in PIDs for the UK is that improving the interconnectedness and reusability of research-related metadata with PIDs will increase the accuracy and completeness of research information, reduce administrative burden for researchers and organisations alike, and enable better decision making and prioritisation of investments in R&I.

Extensive work has already been undertaken to understand the R&I community's needs and priorities and to establish a community coordination and guidance group, the Research Identifier National Coordinating Committee (RINCC). There is a solid evidence base for the nature and scale of potential benefits of PIDs for UK R&I.

The theory of change outlined here sets out the next steps needed to deliver the benefits and gains that have been pinpointed.

With investment in technical and community support for enhanced PID integrations in the information systems which underpin the modern research enterprise, the first benefit would be reduced administrative burden for researchers and managers. As the breadth and depth of information associated with PIDs grows, this will enable improved strategic decision making for institutions, funders, and policy makers. Longer term, improvements to the openness and transparency of the research information ecosystem will ensure that they remain dedicated to serving the evolving needs of the research community, helping to maintain the future resilience and competitiveness of UK R&I.



## Appendix C: Case study - PIDs for projects

Taking research projects as one example, we estimate that there are 50,000 projects active at any given time in the UK HE sector. Research has shown that basic project metadata is entered

into administrative systems at least six times<sup>27</sup>, at a cost of approximately £6 each time.<sup>28</sup> This means the UK is wasting at least £1.8M a year just on simple data entry about projects in universities, all of which could be automated.

If we multiply these costs across the roughly 36,000 grants awarded annually to UK researchers who go on to publish 240,000 articles (not to mention all the books, datasets, software programmes, white papers, policy documents, and other outputs that are generated each year) the waste of time, effort, money, and, worst of all, talent is staggering.

As was demonstrated during the Jisc National Strategy research phase, project PIDs are the least well understood of the priority PIDs<sup>29</sup>. Projects are complex objects, the chief value of which lies in what they contain (other PIDs, metadata, and information about their relationships to each other). Many projects run for a number of years, across multiple institutions, receive funding from a range of sources, employ people on short- and long-term contracts, and produce outputs years after their official close. Currently all of this information is held piecemeal across siloed systems, with much of the context never being captured at all. This creates an unnecessary void, which is slowing decision-making, obfuscating innovation and innovators, and leading to overwhelming administrative burdens among research and support staff alike, as noted above.

At present, the best source of data on projects and related entities comes via expensive, proprietary Current Research Information Systems (CRISs) which enable institutions to track their internal administrative processes (such as human resource functions, logistical requirements, research income) as well as their internal-external interactions (including publications, grant and project management and so forth).

The formalisation of a project PID removes dependency on large (or small, emerging) commercial CRIS suppliers. This is important as the interactions between and across projects (analogously between and across CRISs) are often hidden from the institutions themselves. The sector is currently subject to 'vendor lock-in' where an institutional customer is trapped by a specific vendor's scholarly infrastructure across the breadth of tools and services. This results in a loss of control of and insight into the research enterprise, as well as cutting competition, and slowing dissemination rates. There is no incentive for CRIS vendors to invest in additional functionality which could potentially cannibalise demand for their other products or make it easier for customers to change to another product or service.

Comprehensive adoption of open PIDs for projects would provide a means for sharing data between organisations, enhance strategic analysis of research activity, and ensure data portability and long-term preservation of the research record.

---

<sup>27</sup> Klausen, M.H. (2017) 'Even Minor Integrations Can Deliver Great Value – A Case Study', *Procedia Comput. Sci.*, vol. 106, pp. 153–159, doi: 10.1016/j.procs.2017.03.011.

<sup>28</sup> Research Consulting. (2014) 'Counting the Costs of Open Access', London Higher and SPARC Europe. <http://www.researchconsulting.co.uk/wp-content/uploads/2014/11/Research-Consulting-Counting-the-Costs-of-OA-Final.pdf>

<sup>29</sup> Murphy, Fiona and Jones, Phill (2020) Jisc PID Agency: Focus Group Report. <https://repository.jisc.ac.uk/8166/>