

MORE+BRAINS

Revised cost-benefit analysis for the UK PID Support Network



18 November 2022

DOI: [10.5281/zenodo.7356219](https://doi.org/10.5281/zenodo.7356219)

Photo by Pixabay



1 Background

Persistent identifiers (PIDs) are unique digital codes associated with an entity (such as a person, publication, or organisation) and linked to descriptive information about that entity (metadata). They are a critical component of the UK and global research infrastructure.

In 2021, the MoreBrains Cooperative published a cost benefit analysis for sector-wide PID adoption in the UK¹, which included a proposal to set up a national PID Support Network (PSN). The PSN was further supported by a case for investment², which found that the PSN would save researchers and administrators many thousands of days of work per year that could then be better spent engaged in impactful research.

Since the original report, creation of the PSN has been endorsed by a UK government commissioned independent review of research bureaucracy, led by Professor Adam Tickell in 2022³.

The Review endorses the proposal for a PID consortium made by MoreBrains in their report: 'The case for investment in a UK persistent identifier strategy: Resilience, insight, and leadership in global research and innovation' and would recommend extending this model to other facets of digital research platforms as appropriate.

This update to our original 2021 cost benefit analysis includes a number of methodological improvements.

2 Conclusions

Based on the improved approach, we now estimate that university researchers and administrators waste **55,000 person days** per year unnecessarily rekeying metadata, at a financial cost to the sector of nearly **£19 million**. In the previous version of the analysis, we made no estimate of the amount of time spent rekeying metadata but estimated the financial cost to UK universities to be just over £4.2 million per year.

Considering the cost of implementing the five priority PIDs with support from the PSN, the cost of the PSN itself, and accounting for benefits building over time as adoption increases, we estimate a net saving to the UK higher education sector of over **£45 million** over a five-year period. Our previous estimate was about £5.7 million.

These direct cost savings can be redirected into research itself, with associated downstream economic gains. The total benefit to the UK economy would be £315 million over the five year period.

¹ Brown, Josh, Jones, Phill, Meadows, Alice, Murphy, Fiona, and Clayton, Paul, 'UK PID Consortium: Cost-Benefit Analysis', Zenodo, Jun. 2021. doi: [10.5281/ZENODO.4772627](https://doi.org/10.5281/ZENODO.4772627).

² Brown, Josh, Jones, Phill, Meadows, Alice, and Murphy, Fiona, 'The case for investment in a UK persistent identifier strategy: Resilience, insight, and leadership in global research and innovation', Zenodo, Feb. 2022. doi: [10.5281/ZENODO.6012367](https://doi.org/10.5281/ZENODO.6012367).

³ A. Tickell, 'Independent Review of Research Bureaucracy: final report', Independent Review, Jul. 2022. [Online]. Available:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1094648/independent-review-research-bureaucracy-final-report.pdf (Recomendation 19)

3 Changes to the methodology

For both iterations of this analysis, we based our estimates of the benefits of PID adoption on eliminating metadata rekeying for each of three types of entity (publications, funded grants, and research projects). To give a complete picture of the net benefits of the proposed approach, we accounted for the costs associated with the the creation and operation of the PSN, in combination with a concerted effort to drive the adoption of five priority PIDs (DOIs for funded grants, DOIs for outputs, ORCIDs for people, RAIDs for projects, and ROR identifiers for institutions).

3.1 Revised estimate of rekeying frequency

In the first version of the UK CBA, we assumed that publication, grant, and project metadata are only rekeyed into university systems once. Anecdotally, we believed this was an underestimate based on conversations with researchers and research managers at UK institutions. On the other hand, in the absence of concrete evidence, we chose to use this likely underestimate in order to avoid the risk of artificially inflating benefits.

Following further work, specifically an analysis of the benefits of PID investment in the Australian academic sector⁴, we have increased our estimate of the number of rekeying events, based on a survey of Australian research managers and librarians in which we asked how many times metadata for each of the three entity types is rekeyed. We received 27 responses from 21 unique organisations, representing half of all 42 Australian universities. The results indicated that metadata about publications are rekeyed at least 3.1 times on average, while the average for funded grants is 3.25 rekeying events. We also relied on previous research in Norway which indicated that project metadata is typically rekeyed 6 times⁵. As these estimates are likely to be closer to the real number of rekeying events in the UK, we made use of them in our revised analysis.

3.2 New data for cost per entry

Our estimate of the cost of rekeying publication metadata is based on a Research Consulting analysis⁶, republished in 2021, which shows that publication metadata nominally takes 6.73 minutes to enter into each computer system. The use of length of time per publication as the basis for calculations had the added benefits of allowing a consistent methodology across all three entity types and the generation of an estimate of total time spent rekeying, which was not previously possible.

We updated our method to estimate the cost per minute of rekeying metadata for all three entity types. Rather than assuming indicative salaries of £40k and £60k for administrators and managers respectively, we estimated salaries of £95k and £33k for senior and junior researchers respectively,

⁴ Brown, Josh, Jones, Phill, Meadows, Alice, and Murphy, Fiona, 'Incentives to invest in identifiers: A cost-benefit analysis of persistent identifiers in Australian research systems', Zenodo, Sep. 2022. doi: [10.5281/ZENODO.7100578](https://doi.org/10.5281/ZENODO.7100578).

⁵ M. H. Klausen, 'Even Minor Integrations Can Deliver Great Value – A Case Study', *Procedia Comput. Sci.*, vol. 106, pp. 153–159, 2017, doi: 10.1016/j.procs.2017.03.011

⁶ Research Consulting, 'Counting the Costs of Open Access', London Higher and SPARC Europe, Nov. 2014. Accessed: May 09, 2021. [Online]. Available:

<http://www.researchconsulting.co.uk/wp-content/uploads/2014/11/Research-Consulting-Counting-the-Costs-of-OA-Final.pdf>

and £40k for research administrators. We made this change as we have learned through research and conversations that a greater proportion of administrative work and metadata rekeying is done by researchers than we had previously assumed. An average of the three salary levels, plus a 40% assumed overhead provided the total cost of metadata entry per year.

3.3 Updated costs for the PSN itself

We have also revised the costs for setting up and running the PSN itself. The new estimates are based on a more complete assessment of the requirements for the PSN, which now include the development of the organisational structure, job roles, and descriptions. We used data supplied by Jisc for typical industry standard salaries for the specific roles. In addition, we have included expenses like marketing and communications budgets, travel, and training, but have not costed for HR, legal and finance support directly. Instead we have assumed those support functions would be supplied by a host organisation and have included the standard overhead used by Jisc, which is 28%, as an example to account for that.

4 Revised cost-benefit analysis

Table 1 shows the results of the improved methodology, described in section 3, for calculating the total cost of metadata rekeying into institutional computer systems (time and money) by researchers and research administrators. It shows that nearly 55,000 person days, at a cost of £19 million, are wasted every year at UK institutions through the rekeying of metadata that should be automatically transferred between PID registries, and university and funder systems.

As described in our 2021 cost benefit analysis, PIDs represent a network and are subject to network effects – the full benefits are only realised when everyone adopts them and integrates them into their systems. In the third and fourth lines of table 2, we apply a mathematical model⁷ that accounts for the benefit accrued as the levels of PID adoption increase over a five-year period.

We have made no changes to how we calculated the implementation costs of the five priority PIDs, continuing to assume that six FTEs would be required for the PSN, in addition to other costs like travel, accommodation, and training. Our estimate of the cost of the PSN itself is added to the sector-wide cost of implementation at every university, when supported by the PSN, as shown in table 2. As can be seen, the total net savings accounting for all of the above costs and benefits is over **£45 million** over five years.

Written evidence provided by UKRI to the UK parliament select committee on science and technology suggests that, on average, each £1 spent on research and innovation results in a net gain of £7 to the UK economy⁸. With these cost savings of £45 million being redirected into research itself, the PSN would result in £315 million in total benefit to the UK economy.

⁷ D. Kucharavy and R. De Guio, 'Application of S-shaped curves', *Procedia Engineering*, vol. 9, pp. 559–572, 2011, doi: [10.1016/j.proeng.2011.03.142](https://doi.org/10.1016/j.proeng.2011.03.142).

⁸ N. Lamb, 'Balance and effectiveness of research and innovation spending', House of Commons Science and Technology Committee, Sep. 2019. Accessed: May 10, 2021. [Online]. Available: <https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/1453/145305.htm>

Potential total sector savings (opportunity cost)	Number	# of re-uses	# Authors	# minutes / event	Annual time savings / days	Cost per minute	Annual savings	Five year savings
Publication metadata	236,436	3.10	4.00	6.73	45,359	£1.35 - if Senior Researcher	£26,575,833	£132,879,16
						£0.47 - if Junior Researcher	£9,231,605	£46,158,025
						£0.57 - if Admin	£11,189,824	£55,949,121
						average	£15,665,754	£78,328,770
Grant metadata	36,000	3.25		10.00	2,690	£1.35 - if Senior Researcher	£1,575,877	£7,879,386
						£0.47 - if Junior Researcher	£547,410	£2,737,050
						£0.57 - if Admin	£663,527	£3,317,636
						average	£928,938	£4,644,691
Project description	50,000	6.00		10.00	6,897	£1.35 - if Senior Researcher	£4,040,711	£20,203,555
						£0.47 - if Junior Researcher	£1,403,615	£7,018,077
						£0.57 - if Admin	£1,701,352	£8,506,760
						average	£2,381,893	£11,909,464
Total savings from auto-feed of key metadata via new API links:					54,945		£18,976,585	£94,882,924

Table 1: The total costs in time (~55,000 person days), which corresponds to (~ £19 million per year) of metadata rekeying at UK universities

	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5	Total
Potential savings by not rekeying publication, grant, and project metadata		£18,976,585	£18,976,585	£18,976,585	£18,976,585	£18,976,585	£94,882,924
Adoption target (estimated current level of 18%)	18%	20%	40%	67%	76%	85%	
Percentage of benefits based on 'S-curve'	10.2%	11.1%	26%	71%	85%	93%	
Adjusted savings by not rekeying publication, grant, and project metadata		£2,099,238	£4,996,402	£13,390,246	£16,063,525	£17,688,878	£54,238,289
Costs for the PSN		(£809,894)	(£890,883)	(£979,971)	(£1,077,968)	(£1,185,765)	(£4,944,481)
Sector-wide implementation costs for institutions when supported by the PSN		(£1,541,523)	(£1,622,513)	(£1,711,601)	(£1,809,598)	(£1,917,395)	(£8,602,631)
Sector-wide net savings	£1,930,179	£557,715	£3,373,890	£11,678,644	£14,253,927	£15,771,483	£45,635,658

Table 2: With the revised estimates of administrative cost savings, the PID Support Network will bring over £45 million of direct benefit over five years