Mapping the UK's path to a PID-optimized future

ICENZA

Josh Brown (MoreBrains Cooperative)

DASTELERANCO VENETO



Why does the UK need a PID roadmap now?

...after all, we've been talking about PIDs for a while



> Use of persistent identifiers (PIDs) for scholarly publications (with versioning, for example, in case of revisions), such as DOI (preferable), URN, or Handle.

Strongly recommended additional criteria for all publication venues:

> Support for PIDs for authors (e.g., ORCID), funders, funding programmes and grants, institutions, and other relevant entities.



FI. (meta)data are assigned a <u>globally unique and eternally persistent identifier</u>. II. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.



Recognises Identifiers as a key component of digitisation or science <u>https://doi.org/10.1787/1b06c47c-en</u>



Second draft Persistent Identifier (PID) policy for the European Open Science Cloud (EOSC)

DOI for version 2: https://doi.org/10.5281/zenodo.3780423





Jisc to lead on selecting and promoting a range of unique identifiers, including ORCID, in collaboration with sector leaders with relevant partner organisations.

Funders of research to consider mandating the use of an agreed range of unique identifiers as a condition of grant. (Recommendation 8)

3

Open access to research publications - 2018 Independent advice to the UK government Prof Adam Tickell Background: UK Open access workshop (Oct 2017)

Developing a PID Roadmap



Identifying priorities based on workflows

Workflow gap analysis:

- 1. Started with the real-world workflows and pain points
- 2. Imagined 'ideal' workflows based on gold open access
- 3. What are the priority PIDs?
- 4. What are the points of integration needed?
- 5. How does information flow through the system?

Researcher registers for an ORCID iD	Identifiers used:	 Actions : Researcher visits orcid.org, registers for an iD and adds data to their record using free tools.
Institution adds employment data to ORCID record	Identifiers used:	Actions: • Researcher authorizes employer to read/update record • Employer system pushes job role etc. to ORCID
Researcher applies for funding		
Funder pulls in career and CV data from ORCID record	Identifiers used:	Actions: • Researcher authorizes funder to read/update record • Funder system pulls data in using identifiers
Funding application succeeds		
Funder registers a DOI for the grant	Identifiers used:	Actions: • Funder sends data to grant registry • Funder system includes iDs of grantees in data sent
Funder adds award and review citations to ORCID record	Identifiers used:	Actions: • Funder system sends data to ORCID when award is granted
Researcher submits article		
Publisher collects iD of author(s)	Identifiers used:	Actions: • Author authenticates their iD • Publisher asks to read and update ORCID record
Publisher queries ORCID for funding and employment	Identifiers used:	Actions: • Publisher system resolves identifiers to get data • Author confirms associated funding/employment data
Publisher creates automatic funding acknowledgement	Identifiers used:	Actions: • Publisher system pulls in data from grant registry • Publisher system adds grant DOI to article metadata
Publisher checks if an OA policy applies and any terms	Identifiers used: Oses Funder Registry	Actions: • Publisher system queries grant registry for policy terms
Publisher (re)directs article accordingly	Identifiers used: License Ref?	Actions: • Publisher system matches policy to publishing option(s) • Author notified of options and any APCs and discounts
Publisher sends APC invoice to employer or funder	Identifiers used: Open Funder Registry	Actions: • Publisher system resolves organisation identifier to get address etc.
Open Access article is published		
Institutions and funders harvest	Identifiers used:	Actions: • Auto-update sends citation to funders/employers

MOREBRAINS



Five priority PIDs

Grants



Projects



Persons



Organizations

ROR

Outputs





Sketching the roadmap

The first outline of the roadmap suggested five areas of activity that would help the UK to optimize its use of PIDs:

- A mega-consortium ensure equitable access to PIDs by sharing the burden across the sector and building a multi-PID consortium
- Targeted interventions boost coverage and adoption with high-value integrations with PID infrastructures
- Benefits analysis justify engagement by evidencing and evaluating the impact of PID adoption
- Governance engagement coordinate involvement in PID governance (boards, working groups etc.) to mitigate the risk of depending on a small group of PIDproviding organizations
- Sustainability think about the 'P' in PID... Will these providers be around for the long term?

For more information, see the original 2019 report at: https://repository.jisc.ac.uk/id/eprint/7840



Jisc

Background: UK Open access workshop (Oct 2017)

Our research

ALLE PTANCE

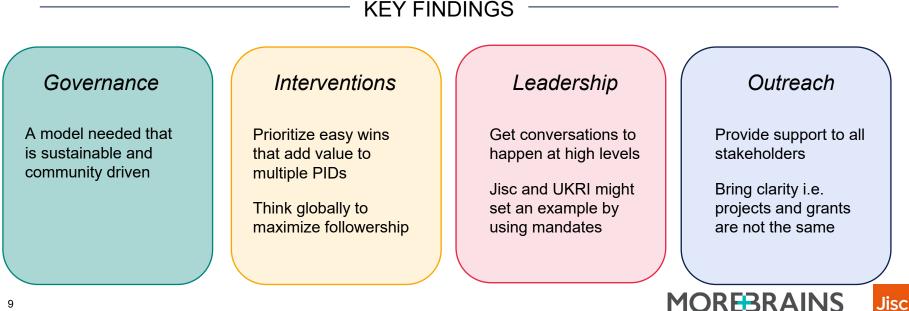
DA CHOICES



+ & LITATIONS

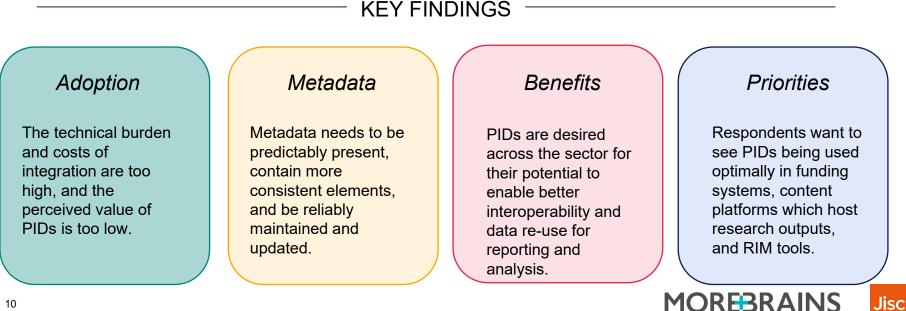
Focus groups

28 attendees across 5 groups 19 different organizations – from publishing trade groups to infrastructure providers to libraries Read more at https://repository.jisc.ac.uk/8166/



Survey

We ran a survey through July and August 2020, which drew 93 responses. 75% from the UK, 70% from research institutions, with the rest drawn from a range of community Read more at https://repository.jisc.ac.uk/8107/ groups.



Background: UK Open access workshop (Oct 2017)

Our approach

11

ALLE PTANCE

DA CHOICES

. & LITARDAS

MOREBRAINS

Jisc₁₁

Bringing the community together: consortium

Adoption and coverage = access and support

Our Consortium Task Group brought together national agencies, funders, institutions, research managers, publishers and PID providers to explore the business case for a 'multi-PID' consortium to lower barriers to PID adoption

ey partners	*	Key activities	*	Value propositions	Critical use cases	Customer relationships	Customer segments
	unders	Construction and and another and another screen another another commission another another another commission another another commission another another commission another another commission another another commission another another another commission another another another commission another another another another commission another another another another commission another another another another another another commission another another another another another another another commission another anoth	Sandaray Albertar Baraspinen Barangan	Efficiency gams policy demands Bener research	Better Sching Exclusion Ex	Customer support team Community group Set5 sensice portalis	Funders
Systems/ platform vendors	utura a	Key resources	2	ingent and puppert	Foctare physics and General Additioners compliance and General Additioners Automation and General Additional Additiona Additional Ad	Channels	Content
	ielated tlatives	PED support hubs	er groon normalise i passa	needs	doctivery metrics autoor expension	Bullionactive community community	services
		RINCC Is	upport	Reduced administration burdent	Credit and Reserving Store States Store States Store S	Crosseef Sponsors Sponsors Vendors	Researchers, Understittensen) Substantensen



MOREBRAINS

Bringing the community together

Assembling a business case for a consortium: the value propositions

"A value proposition is a promise of value to be delivered, communicated, and acknowledged. It is also a belief from the customer about how value (benefit) will be delivered, experienced and acquired."

We took these to the highest level possible to get to the root of the need for PIDs for each group.

Value propositions Efficiency gains Meeting policy demands Better research mgmt and support and serve metadata needs Reduced adminstrativ burden MOREBRAINS



Bringing the community together

Assembling a business case for a consortium: illustrative use cases

We have identified several use cases that cut across groups. These include:

- Enhancing interoperability and data reuse across internal and external systems
- Reductions in administrative burden, cutting overhead and potentially releasing expertise and resources for higher value activities
- Improving data quality, with consequent benefits for evaluation, assessment, analysis and decision support.

Critical use cases Data Reduced nceroperability dminstrative quality burden Better Evaluate activity policy Better data tracking compliance for decision Improved eporting **Business** Better intelligence tracking of Facilitate Discoverability activities of content and policy and careers outputs compliance Content Benchmarking Better Aggregation business metrics intelligence and author discovery experience New/ Credit and Fewer improved recognition distractions tools from research MOREBRAINS



Bringing the community together: consortium

Assembling a business case for a consortium: findings and next steps

"To deliver these efficiencies, we need people with technical and community expertise... providing rigorous and robust technical support for new and existing PID integrations. The value of the consortial approach is that it would provide a framework to deliver this support for every research organisation in the UK, maximising returns on a national investment in PIDs."

"The question is: will the benefits of that adoption justify the costs of the consortium? The answer to this should be provided by a rigorous cost-benefit analysis."

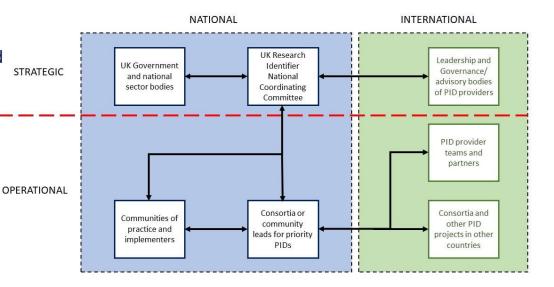


Bringing the community together: governance

Long term strategic engagement

Our Research Identifier National Coordinating Committee will:

- Provide strategic insights to shape the priorities and focus of PID-related activities and policies in the UK
- Liaise with international PID providers and governance structures
- Align PID integrations and developments across sectors to optimise interoperability
- Ensure that all individuals and organisations in the UK research community have equitable access to the PIDs they need



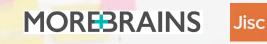


Background: UK Open access workshop (Oct 2017)

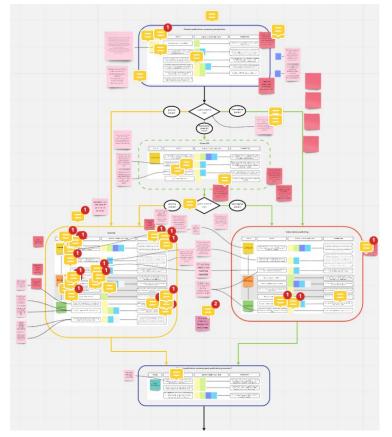
Developing a PID Roadmap: phase 2

DA CHOICES

ALLE PTANCE



PIDs roadmap phase 2 work in progress:



Understanding the perfect PID flow

We are mapping out perfect world PID-enabled workflows for our priority areas

- \circ Funding
- Institutional Research Management
- Content publication (articles and data)

Searching for the perfect point in real-world activities to plug PIDs in and gather metadata

Uncover how every participant in the chain can add value for all the others



PIDs roadmap phase 2 work in progress:

Cost benefit analysis (highlights)

We have pinpointed three areas where benefits accrue: Metadata re-use, automation, and aggregation/analysis.

We are currently collecting data and conducting interviews to enable us to quantify current and potential benefits from the use of PIDs.

- Workflow analysis identify pain points in existing workflows
- Data exchange mapping scale and distribution
- Benefit quantification time and effort savings from PID optimisations
- Modelling different adoption scenarios and timelines
- Community impact analysis identify edge cases, unsupported needs and gaps



PIDs roadmap phase 2 work in progress:

Understanding the pathway to our 'perfect world':

- Take our model workflows and work backwards...
- Identify quick wins
- Locate dependencies and essential integration points
- Create community action plans with justifications for action!

Establishing community oversight with the RINCC

- Setting up the Research Identifier National Coordinating Committee (RINCC) to focus on governance and community accountability
- Establish balanced community representation (different organization types, international partners etc.)
- Facilitate agenda setting and activity design



Stay in touch with the project:

Christopher Brown - christopher.brown@jisc.ac.uk Josh Brown – josh@morebrains.coop

