



PID-optimized research workflows

June 14, 2024 | NTK

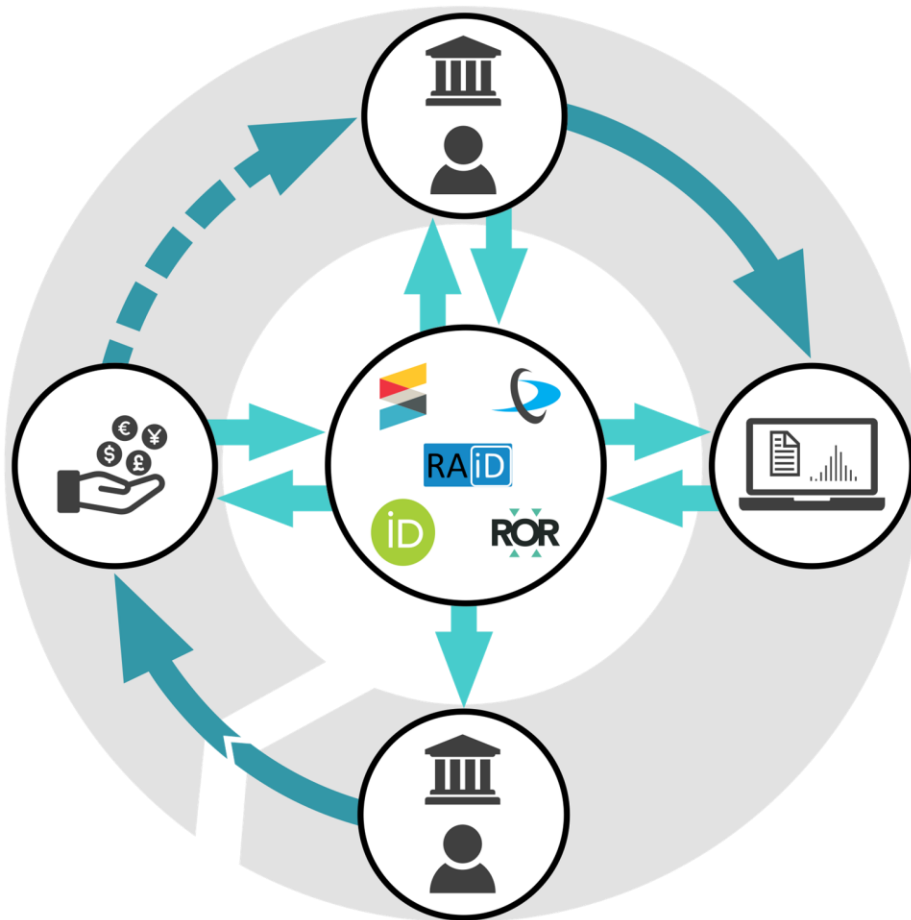
Alice Meadows, Cofounder, MoreBrains Cooperative

MORE+BRAINS

“A PID is a digital identifier that is **globally unique, persistent, machine resolvable**, has an associated **metadata** schema, **identifies** an entity and is frequently used to **disambiguate** between entities.”

US Department of Energy's Office of Science and Technical Information

PID-optimized research cycle



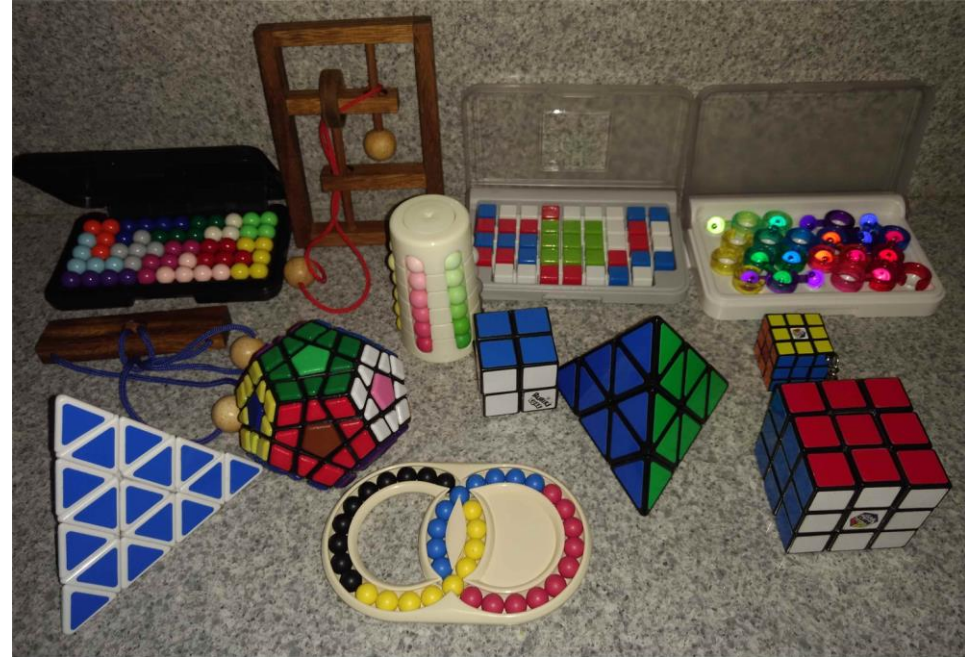
- **Open:** Metadata can move from one system to another
- **Efficient:** Automated movement of information results in complete and timely records
- **Trackable:** Links between PIDs enable connections to be analysed
- **Persistent:** A social rather than technical construct, requiring community/organizational investment by
 - Funders
 - Research institutions
 - Publishers

<https://resources.morebrains.coop/pidcycle/>
<https://doi.org/10.5281/zenodo.4991733>

Brown, Josh, Jones, Phill, Meadows, Alice, & Murphy, Fiona. (2022, September 16). PID-optimised workflows: A vision of a more efficient future. Zenodo.
<https://doi.org/10.5281/zenodo.7085489>

Some examples of research workflow challenges

- Researchers spend too much time on administrative tasks and too little on their research
- Demonstrating ROI in research is challenging - especially over time
- The research evaluation process for grants, and for promotion and tenure, is “detested and a huge administrative burden”
- Research integrity is increasingly under threat



Category:Puzzles - Wikimedia Commons

| PIDs can help us address these challenges

Researchers spend too much time on administrative tasks

- **PIDs allow data to be entered once and reused** across multiple systems, freeing up more time for actual research while also improving accuracy

The research evaluation process is burdensome

- **PIDs can help by storing and maintaining the data** used for evaluation in ORCID records, and pulling it directly from there when needed

Demonstrating ROI in research is challenging - especially over time

- **PIDs enable reliable connections** between researchers, their organizations, grants, and outputs

Research integrity is increasingly under threat

- **PIDs support transparency and trust** through open metadata with clear provenance

Example of a PID solution to a research workflow problem

*“Before using ORCID, an average application used to take a few weeks; formatting took time, getting the publications right took many days of work. If I want to put in an ARC [Australian Research Council] grant now and include all of my research track record, it’s sitting there and ready to reuse and is being continually updated. **This saved me 3-4 days per grant application** - the difference in workload was staggering!”*

Joe Shapter, Pro-Vice-Chancellor, University of Queensland, Australia

<https://aaf.edu.au/wp-content/uploads/JS-Story-FINAL.pdf>



Some examples of research workflow opportunities

- Making data FAIR (**F**indable, **A**ccessible, **I**nteroperable, **R**eusable)
- Better support for collaboration
- Improving attribution and recognition
- Making research more diverse, equitable, and inclusive



Bud - Wikipedia

| PIDs can help us benefit from these opportunities

Making data FAIR

- **You can't be FAIR without PIDs!** PIDs and their metadata underpin many of the FAIR principles

Better support for collaboration

- **PIDs enable reliable connections** between researchers, their organizations, grants, and outputs

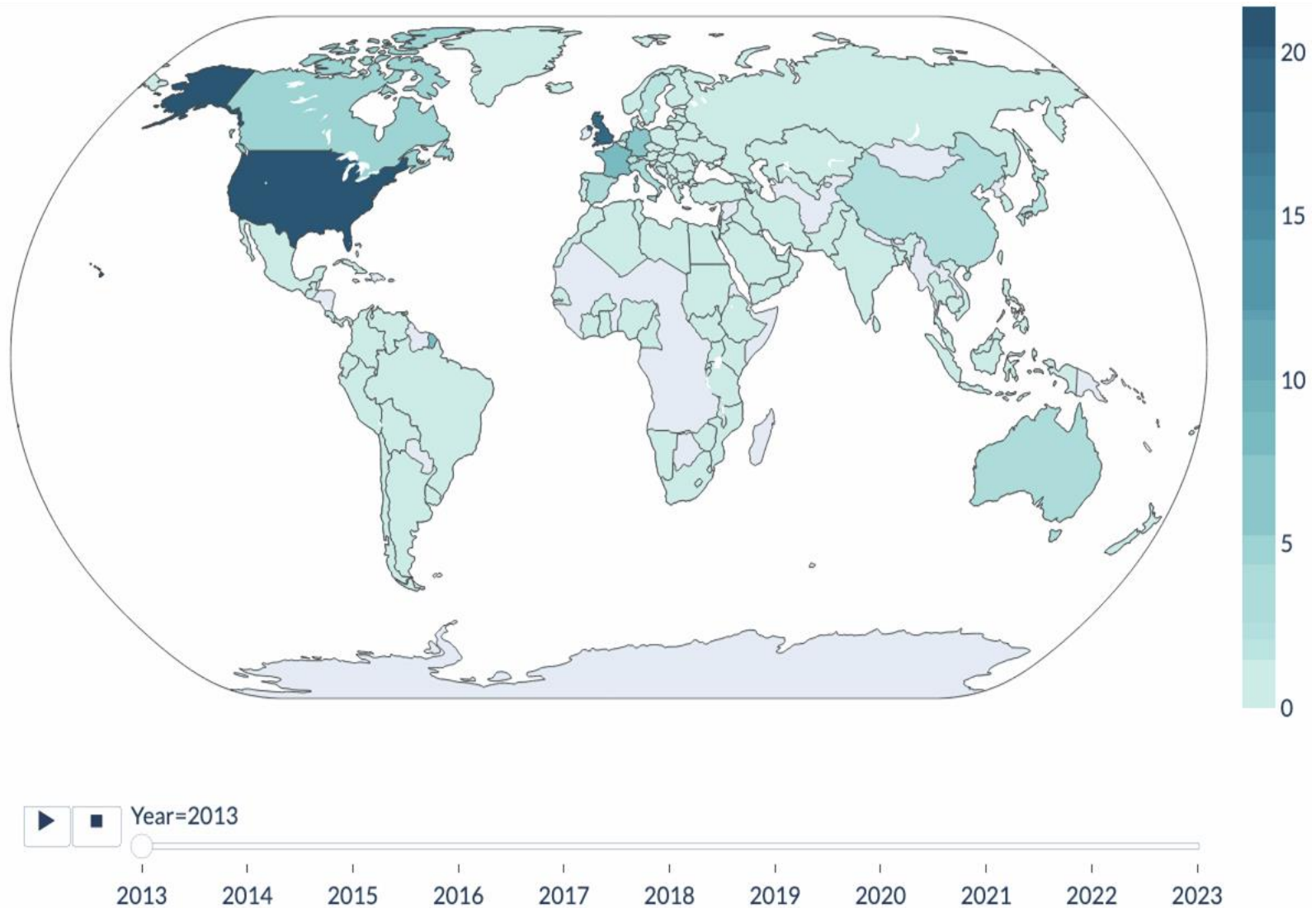
Improving attribution and recognition

- **PIDs help ensure credit** for the right people, organizations, and more (funders, institutions, equipment and instruments, archives, etc)

Making research more diverse, equitable, and inclusive

- **Open PIDs are available to everyone**, for example, through open APIs and metadata

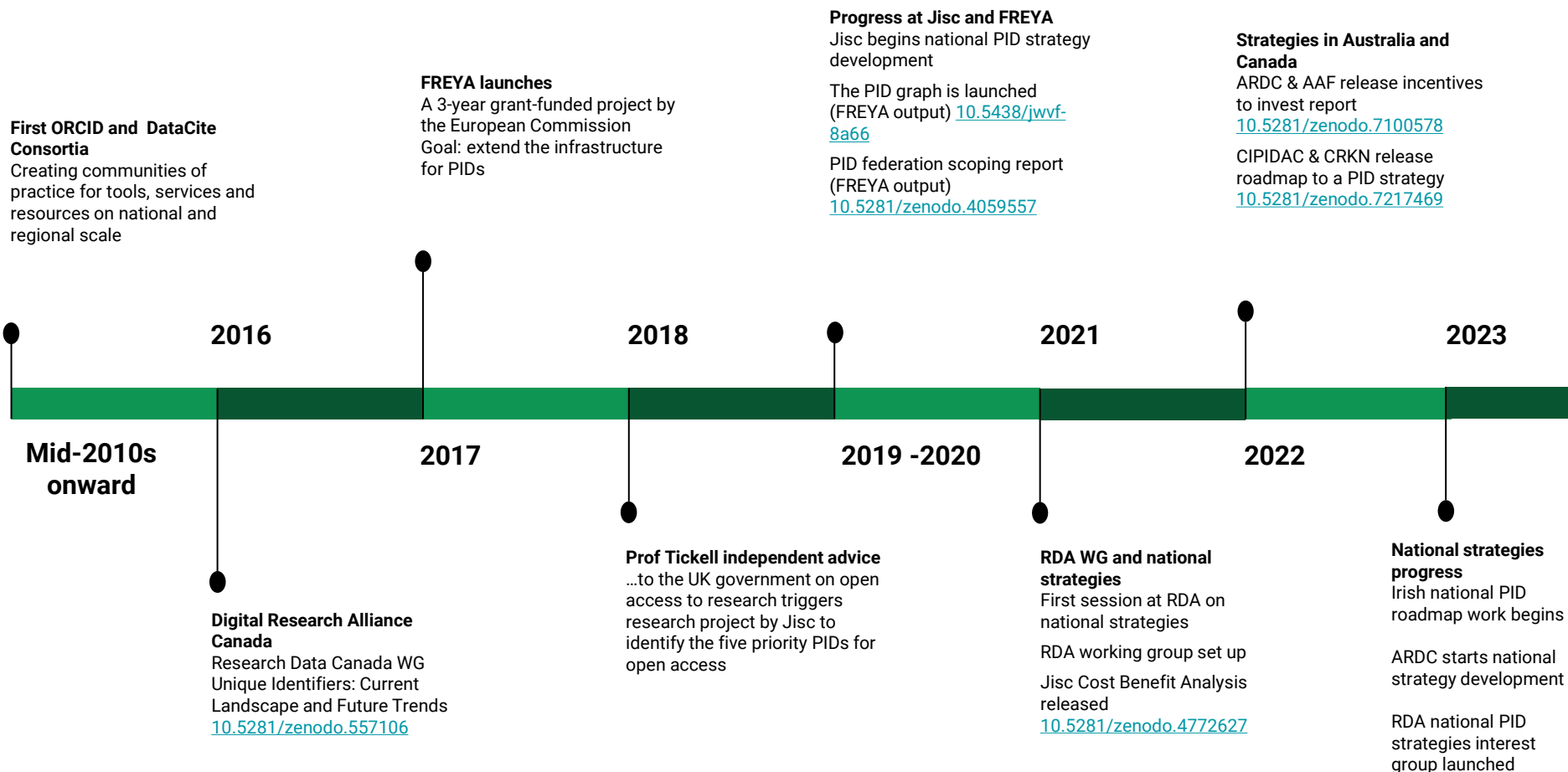
Example of a PID-enabled opportunity



Percentage of global /Irish co-authorships - powered by PIDs!

<https://doi.org/10.7486/DRI.nz80kt123>

National PID strategies are on the rise

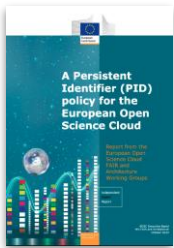


Governmental and inter-governmental interest

G7 Research Compact (2021)

Commitment to work together to improve “...availability, sustainability, usability, interoperability of research data, technologies, infrastructure and services”

<https://policycommons.net/artifacts/1592313/g7-2021-research-compact-pdf-356kb-2-pages/2282082/>



High-level PID policy for EOSC (2020)

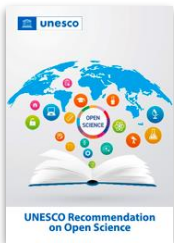
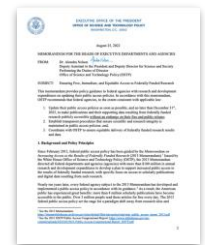
“...a future where PIDs can be used as the preferred method of referring to its assigned entity” <https://doi.org/10.2777/926037>

US Office of Science and Technology Policy memo (2022)

Ensuring free, immediate and equitable access to federally funded research

“...all author and co-author names, affiliations, and sources of funding, referencing digital persistent identifiers”

<https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>



UNESCO Recommendation on Open Science (2021)

“...the definition and attribution of open persistent identifiers as appropriate for each type of digital object, the necessary metadata for their efficient assessment, access, use and re-use...”

<https://doi.org/10.54677/MNMH8546>

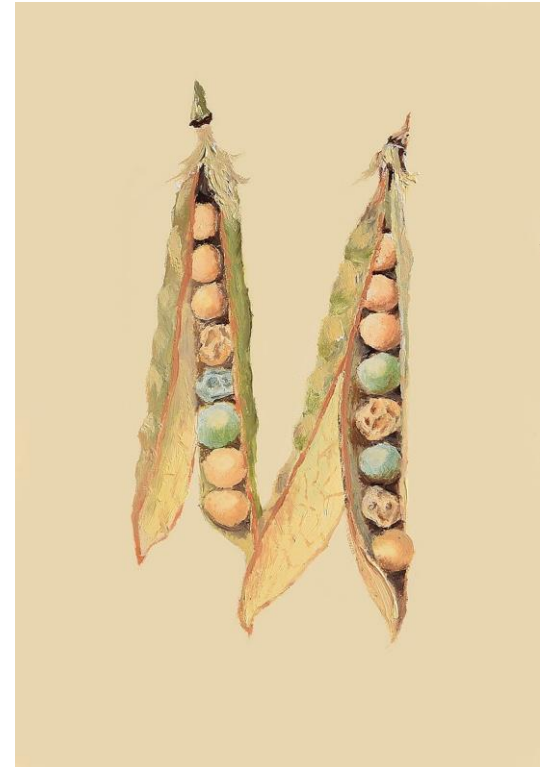
“A PID is a digital identifier that is **globally unique, persistent, machine resolvable**, has an associated **metadata** schema, **identifies** an entity and is frequently used to **disambiguate** between entities.”

US Department of Energy's Office of Science and Technical Information

Desirable PID characteristics

- Stability and persistence
- Global uniqueness
- Resolvability
- Open metadata
- Community governance
- Documentation
- Monitoring and reporting
- Easy assignment and metadata curation
- Interoperability

<https://upstream.force11.org/desirable-characteristics-for-persistent-identifiers/>



File:Fairbanks - Mendel Seed Shape & Color.jpg

Priority entities for PIDs



People: Researchers and other contributors to research (ORCID IDs)



Research outputs: Publications, datasets, and other research contributions (DOIs, eg, Crossref and Datacite)



Research grants: Funding awards (DOIs, eg, Crossref and DataCite)



Research projects: Research project activities, people, and organizations (eg, Research Activity Identifier - RAiD)



Organizations: Funders, institutions, and other research organizations (eg, Research Organization Registry - ROR)

| PIDs for research outputs

- **Digital Object Identifier (DOI)** – PIDs for any physical, digital, or abstract object
 - DOI Foundation – manages DOI Registration Agencies
- **Crossref**
 - ~160M DOIs registered by 17k+ members
 - 1bn+ metadata searches each month
 - DOIs for books, chapters, articles, images, peer review reports, working papers, and more
- **DataCite**
 - 50M+ DOIs registered by 3.2k repositories, 1.4k other organizations
 - DOIs for datasets, physical samples, audiovisual, images, and more



| PIDs for research grants

- **Crossref**

- 122k+ registered to date
- NOT the same as funders' own grant identifiers
- DOIs can be registered for any type of support provided to a research group or individual, eg, awards, use of facilities, sponsorship, training, salary, etc

- **DataCite**

- Coming soon!



MORE+BRAINS

| PIDs for research projects

- Research Activity Identifier (RAiD) — a “container” PID for information about project personnel, organizations, inputs, and outputs
- Developed, maintained, and operated by the Australian Research Data Commons (ARDC)
- Formalized as an ISO standard (ISO 23527:2022)
- ARDC is partnering with DataCite to deliver the RAiD service

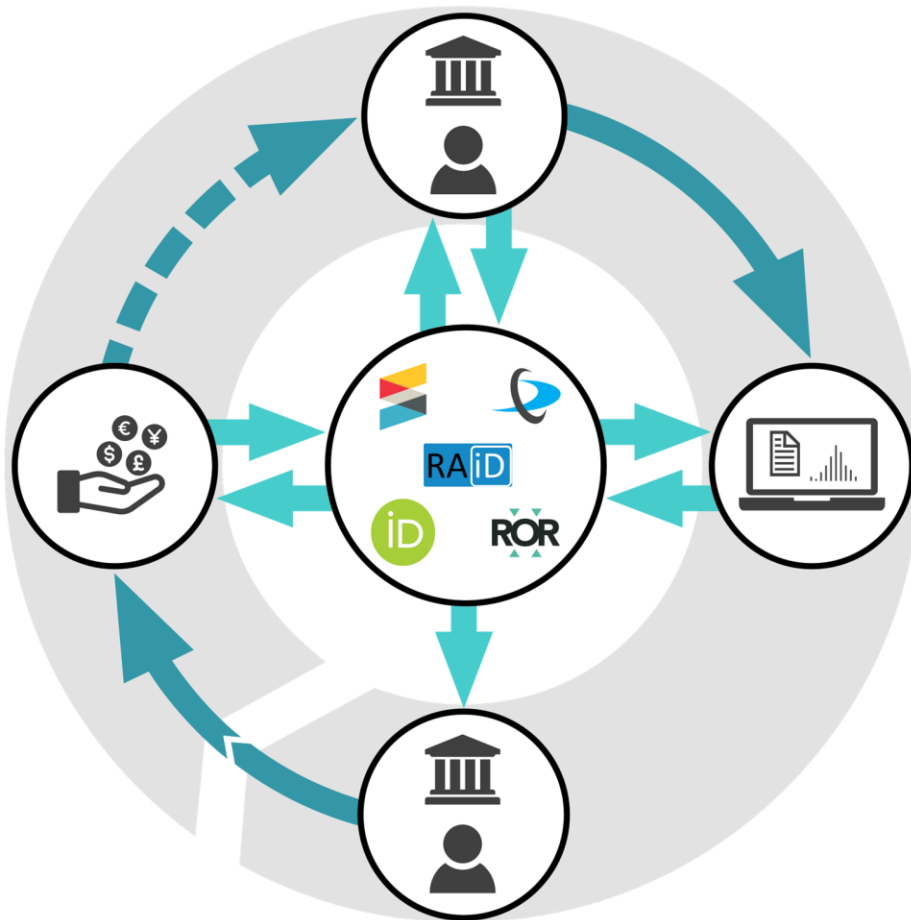


| PIDs for research organizations

- Research Organization Identifier (ROR) — focused specifically on research organization affiliations
- Managed by California Digital Library, Crossref, and DataCite
- 107k+ organizations registered
- Intended for all organizations involved in research:
 - **Includes 35k research funders**
- Will merge with Crossref Open Funder Registry next year



PID-optimized research cycle



- **Open:** Metadata can move from one system to another
- **Efficient:** Automated movement of information results in complete and timely records
- **Trackable:** Links between PIDs enable connections to be analysed
- **Persistent:** A social rather than technical construct, requiring community/organizational investment by
 - Funders
 - Research institutions
 - Publishers

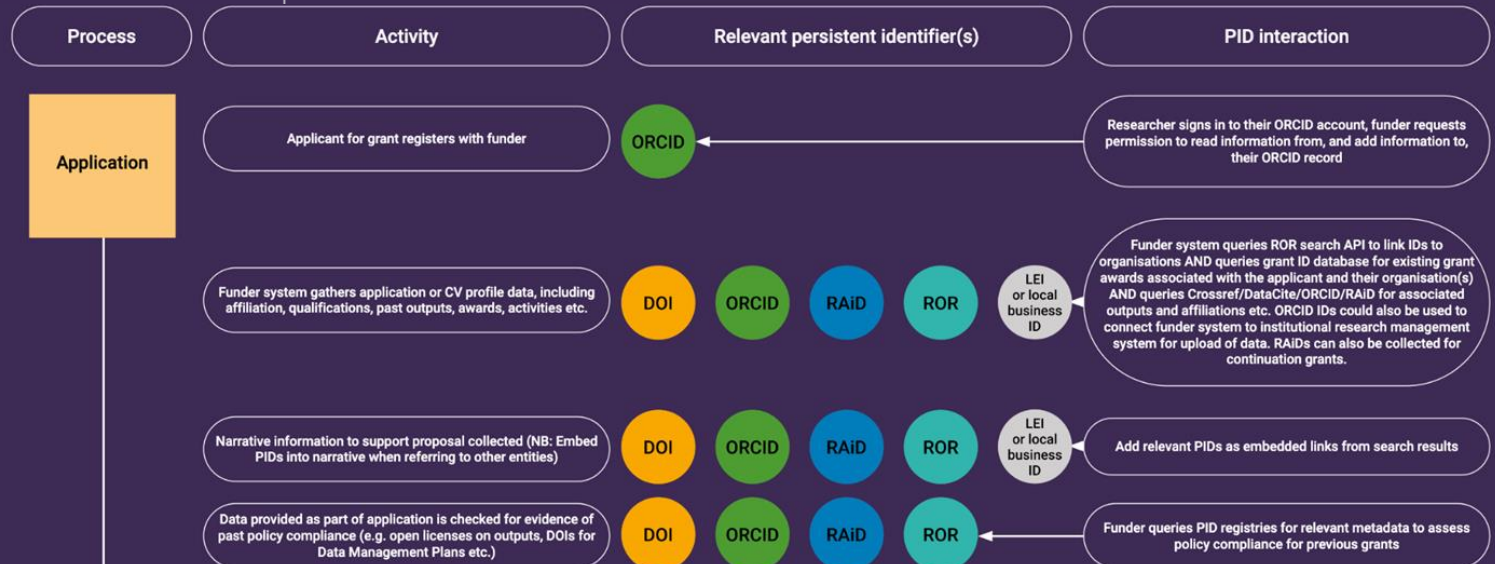
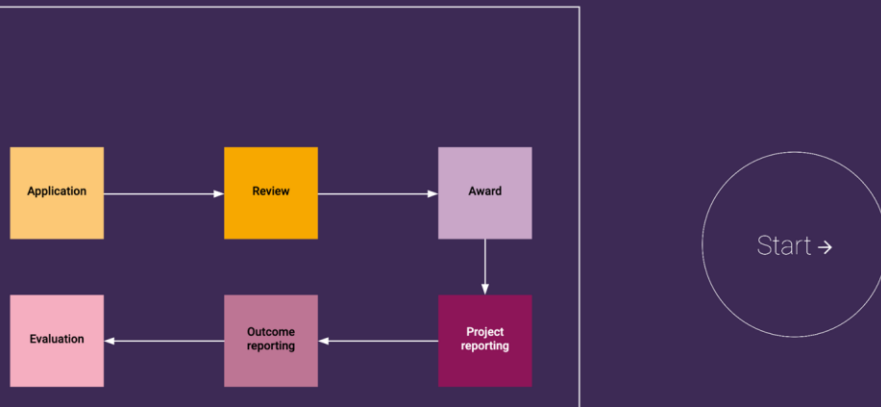
<https://resources.morebrains.coop/pidcycle/>
<https://doi.org/10.5281/zenodo.4991733>

Brown, Josh, Jones, Phill, Meadows, Alice, & Murphy, Fiona. (2022, September 16). PID-optimised workflows: A vision of a more efficient future. Zenodo.
<https://doi.org/10.5281/zenodo.7085489>

PIDs in funding workflows

<https://resources.morebrains.coop/jisc-workflows/jisc-funding-workflow/>

Funding workflow



Three things funders can do to support PIDS

1. Register, maintain, and use a ROR ID for your own organization
1. Register and use PIDs for the grants you award
1. Encourage your awardees to get and use an ORCID ID



File:Numbers 1-2-3.jpg - Wikimedia Commons

Chen, X., Cousijn, H., Hendricks, G., Sadler, S., & Stathis, K. (2023). Guide for funders to support FAIR workflows & enable research tracking. Zenodo.

<https://doi.org/10.5281/zenodo.8289142>