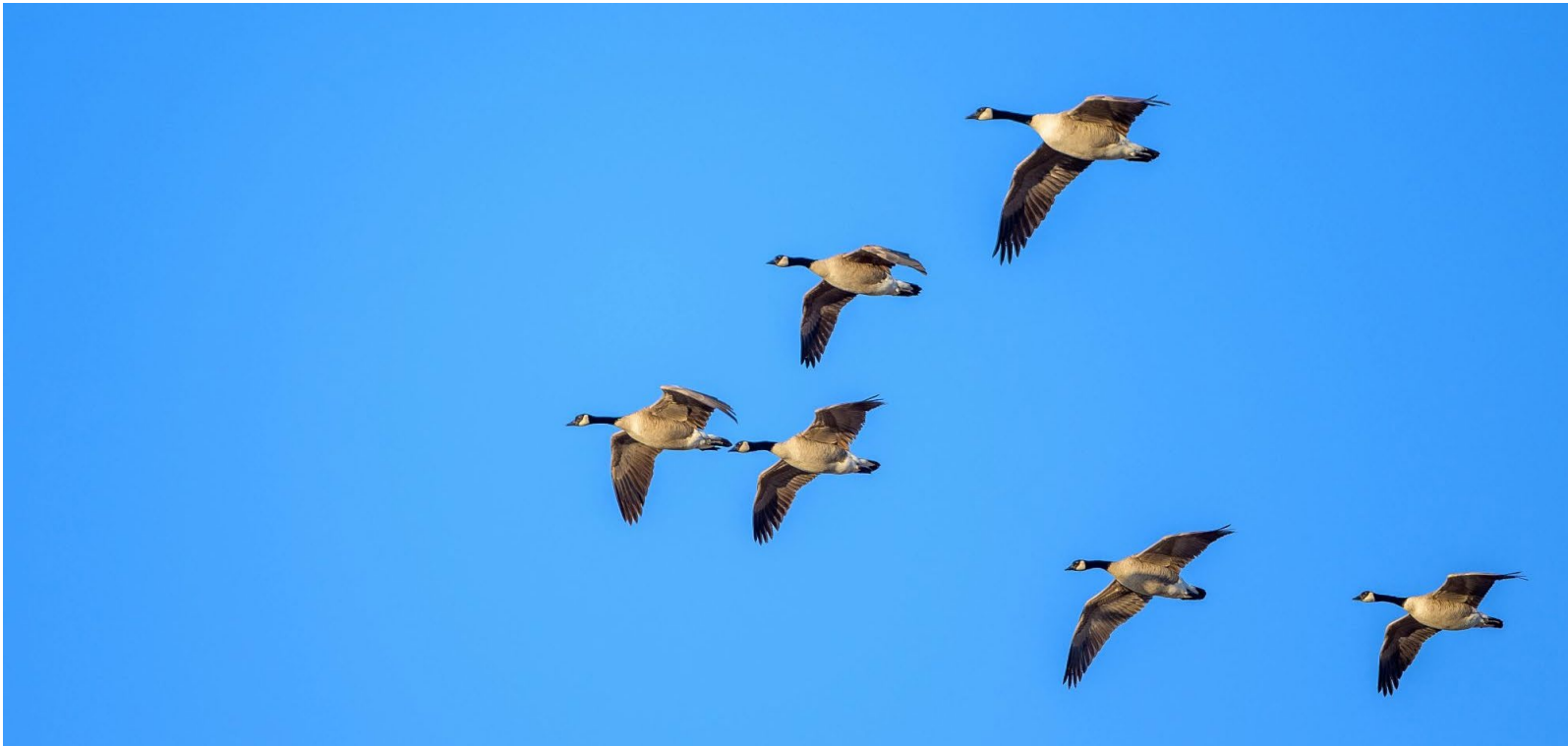


MORE+BRAINS

Investing in an efficient and transparent research ecosystem in Canada

Evidence, analysis, and recommendations for
a Canadian persistent identifier roadmap



La version française suit

7th June 2024

Acknowledgements

This report was commissioned by the Digital Research Alliance of Canada (the Alliance) and the Canadian Research Knowledge Network (CRKN), with funding provided by the Alliance and with substantial guidance provided by the Canadian Persistent Identifier Advisory Committee (CPIDAC).



**Digital Research
Alliance** of Canada

**Alliance de recherche
numérique** du Canada



**Canadian Research
Knowledge Network**

**Réseau canadien
de documentation
pour la recherche**

The authors would like to thank the project partners, the members of the CPIDAC and all the Canadian research community members who so kindly volunteered their time and insight for this project.

This report and its underlying materials are provided under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>)

Executive Summary

This phase of the Canadian PID roadmap project built on the two previous phases to develop a vision, mission, and principles for the strategy; a set of priority PIDs and integrations; and a set of actionable recommendations to deliver the strategy.

The principles highlight the community's need for a strategy that: prioritizes open systems; is inclusive and equitable; embraces community engagement and collaboration; is international in scope; will bring about effective, tangible change; and will support high-quality, trustworthy, and verifiable data.

Community consultations identified eight entities as candidates for assessment and prioritization in the next phase of national PID strategy development (in alphabetical order): data management plans (DMPs); facilities, instruments, equipment; grants; organizations; outputs; people; projects; and software.

The project developed a PID selection matrix, intended to enable easy comparison of candidate PIDs for these entities, in order to support informed decision-making both now and in the future and with criteria that incorporate best practices globally as well as the Canadian community priorities reflected in the vision, mission and principles.

While the Canadian research community will ultimately decide which PIDs to adopt, we have assessed likely candidates for the priority entities. Some are already mature and widely adopted (DOIs for DMPs, grants, and outputs; ORCID for people). However, there is more than one candidate PID for organizations. DOIs for instruments are available but not yet widely adopted. There are no clear PID options for facilities, equipment and software. There is one likely candidate PID for projects, the Research Activity Identifier, which is under active development.

Based on existing work from previous phases of the project, consortial outreach and surveys, and CPIDAC discussions, we conducted a gap analysis, validated by community interviews and focus groups, to identify the components of the technical, social, and policy landscape that need to be aligned to deliver the strategy. Four themes emerged from the gap analysis —governance, technical capacity, leadership and clarity, and community engagement and outreach — and our recommendations are grouped around these themes.

Clear high-level political leadership is needed to implement the PID strategy, to help unlock resources and to bring PID integrations to the top of organizational priorities. Oversight of the strategy development, especially with regard to community inclusion and representation and the selection of PIDs on which to focus, will be critical to community buy-in. There is limited technical capacity for PID integration, and we offer detailed recommendations for each section of the community for first steps or priorities. Finally, we emphasize the need for targeted outreach and consistent messaging for key communities and for different groups of practitioners in the space.

Table of Contents

1. Background	4
2. Synthesizing the evidence	4
3. Findings and outputs	6
3.1 Vision, mission, and principles	6
3.1.1 Vision	6
3.1.2 Mission	6
3.1.3 Principles	6
3.2 Identifying priority entities and candidate PIDs	7
3.2.1 Prioritization workshop outcomes	8
3.2.2 Priority entities	9
3.2.3 PID awareness and availability	10
3.2.4 The PID selection matrix	10
3.2.5 PIDs for priority entities	12
3.3 Landscape analysis	14
3.3.1 Gap analysis	14
3.3.2 Institutions	16
3.3.3 Funders and agencies	17
3.3.4 Key policy points	19
4. Recommendations	20
4.1 Leadership and clarity	20
4.2 Governance	20
4.3 Technical capacity	21
4.4 Community engagement and outreach	22

1. Background

Starting in 2021, the MoreBrains Cooperative team have been working with the Canadian Research Knowledge Network (CRKN)¹ to help lay the foundations for a national persistent identifier (PID) strategy in Canada. This work is being funded by the Digital Research Alliance of Canada ('the Alliance')², and is now in its third phase. The first phase was largely consultative, taking in existing work, collaborations, consortia, and policies to identify steps that needed to be taken to create a fully-fledged strategy. Based on our findings, in the second phase we developed a mix of communications outputs, including briefings for key stakeholders, introductory materials for those new to PIDs, and analyses of specific community needs and challenges around PID adoption and integration. This third phase has focused on generating specific and concrete actions and priorities for PID integrations, informed by the consultations and by our research from previous stages. Deliverables include a prioritized list of PIDs, system integrations, and capacity-building activities.

We have taken a three-step approach to this work, aiming to deliver a clear, evidence-based set of high-priority actions that are focused on specific challenges, and that identify a clear ask for each group of stakeholders:

1. Working closely with the project team and the Canadian Persistent Identifier Advisory Committee (CPIDAC) to develop a vision, mission, and set of principles for the strategy, as well as conducting a gap analysis to evaluate current knowledge and sector capacity.
2. Identifying priority PIDs and integrations, using the draft 'PID selection matrix' from the second stage of this project, together with the results of the gap analysis and supplemented by additional landscape research.
3. Conducting six interviews with technical staff at funding organizations and two focus groups with institutional representatives, calibrated to provide specific, actionable insights for an ongoing programme of activity to deliver the strategic vision.

2. Synthesizing the evidence

Over the three project phases, we have built a large body of evidence, and much of the work described in this report has focused on processing this information so that it can be used for decision-making. This has resulted in four separate pieces of synthesis which shape the recommendations provided here:

1. A vision, mission, and principles

The vision and mission articulate the ultimate goals of the PID strategy, while the principles will inform the design and delivery of the strategy, including informing decisions and prioritization.

The evidence base for these statements came from the interviews, focus groups, workshops, surveys, and other community consultation activities undertaken during previous project phases and by the Canadian ORCID and DataCite consortia. Discussions with the CPIDAC about community concerns and ambitions during phases one and two were also included in the analysis. Revisiting contemporaneous notes and the summaries presented in our reports gave us a clear sense of themes that mattered to the grassroots research community members we spoke to, and the vision, mission, and principles went through several rounds of iteration with the project team and the CPIDAC before being finalized. For more detail, please see section 3.1 below.

¹ <https://www.crkn-rcdr.ca/>

² <https://alliancecan.ca/en>

2. A research information lifecycle diagram

Understanding the practical issues that PIDs and their associated metadata could solve in a research workflow context is an important step in this process. In phase two of the project, we held two virtual workshops that approached issues in research information from different perspectives, but with a common focus on information loss or friction in the exchange of data.

The findings from these workshops were then structured as a research information lifecycle workflow diagram, breaking out the activities which typically occur at various stages from the conception of research ideas through to post-project reporting, evaluation, and impact assessment.

Presenting the challenges in this way enabled us to overlay community suggestions for steps that could be improved with PID integrations, and to map consistent or highly impactful interventions to both the community and the activity that would benefit from them. It also enabled us to pinpoint the entities for which information was needed at each step. These entities were then taken forward as a priority for further assessment in terms of PID availability and suitability.

The key findings about pain points and priority entities are set out in section 3.2.2.

3. A PID selection matrix

To help evaluate the PIDs available for the priority entities, we needed consistent and transparent criteria, which take account of the goals and values set out in the vision, mission, and principles, and which incorporate community priorities and concerns identified during the project to date. The raw outputs from eight workshops were re-analyzed to identify features of PIDs that reflect community priorities, and we also revisited previous syntheses from our end-of-phase reports or specific project outputs, such as the research information lifecycle workflow diagram.

The criteria also needed to take account of established good practices in PID provision, in terms of both technology and governance. We relied for these on work undertaken by previous PID-related projects, such as the European Commission-funded ORCID and DataCite Interoperability Network (ODIN) project³, the European Open Science Cloud PID Policy⁴, and the Principles of Open Scholarly Infrastructure⁵.

Based on all these inputs, nine categories, comprising 39 separate criteria, were developed to inform the comparison and selection of PIDs to be included in the strategy. The details of this process can be found in section 3.2.4, and a full list of the inputs used for the creation of the matrix are provided in the first tab of the PID selection matrix spreadsheet.

4. A gap analysis

Understanding the assets, capabilities, and systems that are already available to support the delivery of the PID strategy, and setting them alongside known gaps was also critical to our synthesis. It also enabled us to infer additional components or steps that need to be taken to fill gaps, pinpoint dependencies, and shape concrete recommendations for actions and interventions.

Every project activity and output to date was fed into the gap analysis, resulting in a comprehensive, high-level overview of the landscape that the PID strategy will operate in. The analysis was mapped out graphically using the virtual whiteboard tool Miro, and the findings were then categorized and presented as a slide deck. The findings are set out in section 3.3.1 below.

³ <https://cordis.europa.eu/project/id/312788>

⁴ <https://doi.org/10.2777/926037>

⁵ <https://openscholarlyinfrastructure.org/>

3. Findings and outputs

This report is based on our findings from all three phases of the project to date. The consultations and research conducted in phases one and two were synthesized into focused outputs, each of which is described in the relevant section below. The first output in this phase of the project was the vision, mission, and principles statement, which underpins the whole strategy and which, combined with the priority entities workflow and analysis from phase two, shaped the design of the PID selection matrix and gap analysis. These, in turn, informed the research questions we used for the interviews and focus groups that we conducted as part of the landscape analysis.

The recommendations that resulted from this work have been mapped back to the vision, mission, and principles in order to support future decision-making and ensure the clear strategic focus needed for the success of the project.

3.1 Vision, mission, and principles

The vision, mission, and guiding principles for the Canadian national PID strategy were based on our findings from the community consultations carried out in previous phases of the project, and developed in close consultation with the CRKN and Alliance project team. A draft version was then shared with the CPIDAC ahead of their in-person meeting in Ottawa on March 5, 2024, where Alice Meadows of MoreBrains conducted a workshop to gather additional input and feedback from the group. This resulted in some changes to the vision and mission, edits to one of the principles, and the addition of a new sixth principle. The final versions are as follows:

3.1.1 Vision

Comprehensive persistent identifier (PID) adoption will underpin an effective and equitable digital research information landscape in Canada, enabling open and persistent connections between research inputs, activities, outputs, and outcomes.

3.1.2 Mission

Recognizing that persistent identifiers are a foundation of a modern digital research system, the national PID strategy will provide a framework to support the widespread adoption of persistent identifiers across the whole Canadian research ecosystem. It will reflect the priorities of all research communities in Canada, recognizing that research takes place within and across different cultures, disciplines, and languages. It will be responsive, equitable, and inclusive, with transparent and consultative processes at its core. The strategy will be built on evidence — from analyzing community needs to evaluating its effectiveness — adapting and evolving over time to ensure technical, organizational, and social sustainability. It will be nationally focused but internationally networked, incorporating best practice wherever it may be found in order to support the use of PIDs to uniquely identify a range of entities across the research lifecycle; support interoperability between systems and information resources; and improve the flow of accurate (meta)data.

3.1.3 Principles

These six principles have been developed in response to community consultation undertaken during 2022 and 2023. They are designed to act as a guide to the development and execution of the Canadian national PID strategy, to help to shape decision-making, and to keep the strategy focused on community needs and priorities as it evolves in the future.

Principle 1: Open systems will be prioritized wherever feasible

Part of what makes PIDs so valuable is their ability to connect information in different systems and between organizations. Using PIDs, metadata, and registries as bridges between information systems requires maximum access to these tools and minimal limits on data reuse. We will therefore prioritize

systems that provide open Application Programming Interfaces (APIs), permissive and clear licensing for data access and use, and that have community governance to assure openness in the future.

Principle 2: The strategy, the activities it shapes, and their outcomes must be inclusive and equitable

The benefits of the PID strategy must reach every community engaged in research in Canada.

Acknowledging that the value of PIDs grows the more they are used, we will strive to ensure that the value generated by PID adoption is not concentrated on those groups and organizations that already enjoy relative advantages of wealth and status. In practice, this means identifying less advantaged stakeholders, and actively working to secure resources, providing additional support or services, and creating accessible toolkits for those who need them.

Principle 3: Community engagement and collaboration will be a central component of the strategy

Our strategy has emerged from community consultations and research. We will ensure that this community-centric approach continues to evolve, proactively fostering collaboration and facilitation across stakeholder groups, including those outside of the existing PID community, in order to understand and amplify their needs. We will raise awareness of PIDs and their benefits among all those contributing to research in Canada through outreach efforts designed to engage, explain, and educate.

Principle 4: The national strategy will be international in scope

We recognize that Canadian research activity does not only take place within our borders. Canadian researchers work with collaborators around the globe, receive funding from bodies based outside Canada, and publish in venues managed in other countries. We will be an active voice for our nation among this international network of partnerships, and will prioritize internationally used PIDs, schemas, and infrastructures to ensure that national and international systems can interoperate and share information about research activities, wherever they take place.

Principle 5: The goal of the strategy is effective, tangible change

The adoption and integration of PIDs in the tools and systems used every day by researchers and those who support them requires investment and adaptation. These tools and systems must deliver real value, freeing researchers to be more effective by improving discovery, making their working lives easier, and reducing administrative effort. In concrete terms, every intervention proposed in the strategy will be targeted at a specific challenge, or opportunity with clear actionable goals and community-validated criteria for evaluating its impact.

Principle 6: The strategy will support high-quality, trustworthy, and verifiable data

High-quality, verifiable data and metadata enables more insightful analysis, improves trust through reliability, and anchors data to its creators, increasing the interoperability and integrity of information across systems. The PID strategy will prioritize systems that foster trust and enhance the quality of PIDs and their metadata, while respecting the security and privacy of individual researchers and contributors.

3.2 Identifying priority entities and candidate PIDs

The report *Towards a national PID strategy for Canada - Vers une stratégie nationale sur les PID pour le Canada*⁶ outlines the findings of MoreBrains⁷ investigation of the current PID landscape in Canada, and community expectations for what a PID strategy must deliver for them. The analysis set out in this section addresses parts of recommendations 3 and 4:

⁶ Brown, Josh, Jones, Phil, Meadows, Alice, & Murphy, Fiona. (2022). Towards a national PID strategy for Canada - Vers une stratégie nationale sur les PID pour le Canada. Zenodo. <https://doi.org/10.5281/zenodo.7217469>

⁷ <https://www.morebrains.coop/>

3. Research community needs and barriers (e.g. to understand the motivations for non-engagement)

4. Identify a core set of community goals with associated PIDs and integration needs

In April and May 2023, in partnership with the project coordinators from CRKN and the Alliance, MoreBrains convened two virtual community workshops. The first focused on information exchange between research-performing organizations (such as universities and research institutes) and research funders; the second focused on the integration and uptake of PIDs within institutions.

3.2.1 Prioritization workshop outcomes

The findings of the first workshop are presented below and in section 3.2.2. They detail respectively the challenges and pain points flagged by participants; and the entities for which comprehensive PID adoption would be most helpful in addressing those challenges. Participants included representatives from funders, including the Tri-agency group and the Fonds de Recherche du Québec, and representatives from six universities.

Section 3.2.3 includes our findings from the second workshop, and sets out some preliminary pointers for additional research and indications of the current state of PID adoption and integration in Canada. The workshop was attended by a range of product, systems, and analysis professionals from funders, service or platform providers, and universities,

Challenges and pain points

Issues flagged from inception to funding application

Preparing funding application:

- Assembling and structuring CV data
- Entering similar information in multiple forms
- Difficulties in reuse of information across grant applications

Internal assessment

- Ad hoc or non-standard approaches to sharing application data
- Challenges identifying appropriate expertise and avoiding conflicts of interest

Application/submission

- Repetitive, manual data entry produces errors and wastes researcher time
- Complex grant applications have multiple information exchange instances with the potential for information loss
- Variety and variations in data collected
- Limited space to list outputs, collaborations, etc.

Peer review

- Identifying potential conflicts of interest
- Information about peer review activity is not visible
- Over-reliance on a small pool of reviewers

Issues flagged from funding through to project execution

Formal award and acceptance

- Ambiguous and inaccurate records of organizations and people
- Grant or award numbers not guaranteed to be unique and with no formally associated descriptive information
- Duplicating data entry and mismatched records at different organizations

Project set-up

- Manual data entry with little interoperability or reusability
- Difficult to track entities across systems or share information
- Challenges in tracking participation and crediting contributions
- Ad hoc or non-standard approaches to sharing project data

Project monitoring

- Outputs without PIDs are difficult to track over time
- Inconsistent and incomplete acknowledgements
- Proprietary, incomplete, or speculative information
- Duplicating data entry and mismatched records at different organizations

Issues in reporting and evaluation

Output/outcome tracking

- Difficult to check outputs for policy compliance
- Variable naming and description conventions make data hard to use
- Capturing data from partners is difficult, manual work
- Hard to understand career trajectories and impacts

Reporting and compliance

- Qualitative data time-consuming to assemble
- Inconsistent and incomplete acknowledgements
- Retrospectively reconstructing information about what has been published is challenging

Impact and evaluation

- Poor data exchange and interoperability between systems and organizations
- Dependent on third-party (often proprietary, costly) aggregators/databases
- Hard to understand career trajectories and impacts
- Difficult to track social or economic impacts back to project activities/outputs

3.2.2 Priority entities

While some entities can be identified from the PID interventions proposed by workshop participants, others are implicit in the problems raised or solutions proposed. For example, ‘reuse data from PID registries’ was suggested as a solution to multiple examples of difficulties with data collection and reuse. In these cases, we inferred the most relevant PIDs by analyzing the context (such as project management or grant application).

Having collated all the entities referenced or implied across the workflow into a single, deduplicated list, the following eight entities (in alphabetical order) are candidates for assessment and prioritization in the next phase of national PID strategy development:

- DMPs and similar records
- Facilities, instruments, equipment
- Grants
- Organizations (institutions, funders, etc.)
- Outputs (data, articles, books, etc, as well as “non-traditional”)
- People
- Projects
- Software

As well as establishing community priorities for these entities, we will need to analyze available PID options and their maturity, and take this into account when deciding on the core set of PIDs for the national strategy. For example, the Open Researcher and Contributor Identifier (ORCID)⁸ is a well-established, *de facto* global standard for researcher identification, with a national consortium and engaged community in Canada. Conversely, identifiers for software and projects are at a much earlier stage of development in terms of infrastructure, governance, and community.

3.2.3 PID awareness and availability

The second workshop shed some light on current levels of awareness of commonly used PIDs in research-performing organizations, funders, and platform/systems providers.

Across all three groups, ORCID IDs for people, Research Organization Registry⁹ (ROR) IDs for organizations, and Digital Object Identifiers¹⁰ (DOIs) for outputs (especially journal articles and datasets) are well known. In the platform/systems provider group, these were the only PIDs that were discussed, apart from a single mention of the International Standard Serials Number¹¹ (ISSN) for journals. Funders also mentioned the Research Activity ID¹² (RAiD) for projects, as well as identifiers for instruments and Research Resource IDs¹³ (RRIDs). Institutional representatives also mentioned these PIDs, along with some other identifiers for organizations and people (such as the proprietary Ringgold ID¹⁴ or Scopus ID¹⁵).

It is clear that in Canada, as in many other countries, there is limited awareness of the availability, maturity, and status of PIDs for many of the entities listed in section 3.2.2. This is significant because currently there is no widely adopted PID for some entities (for example, projects), and yet they are absolutely critical in order to address many of the workflow issues raised in the workshop.

3.2.4 The PID selection matrix

For several priority entities, there is more than one candidate PID, each with a variety of characteristics that differentiate them. For example, PIDs for researchers include both proprietary PIDs, such as the Scopus Author ID, and open PIDs, like ORCID or ISNI¹⁶. Of the open PIDs, ISNIs can be used for anything with a name (including, for example, fictional characters), whereas ORCID is focused on the research sector.

A consistent and coordinated approach to national PID adoption is needed to enable coverage to reach a level where the benefits of PIDs are reliably delivered to the widest possible range of stakeholders. Consistency also brings other benefits, such as familiarity with the feature set of a given PID, enabling the community to make the best use of these PIDs in research information systems.

We have therefore developed a PID selection matrix to enable easy comparison of PIDs, based on a clear understanding of their characteristics and unique features. Using this matrix will support informed decision-making on the selection of a subset of the available candidate PIDs, in order to make the best use of limited resources. Taking a structured approach to these comparisons will enable them to be revisited, adapted, or updated; it will also help to assess new PIDs and/or additional

⁸ <https://orcid.org>

⁹ <https://ror.org>

¹⁰ <https://www.doi.org/>

¹¹ <https://issn.org>

¹² <https://raid.org>

¹³ <https://scicrunch.org/resources>

¹⁴ <https://www.ringgold.com/>

¹⁵ <https://www.scopus.com/freelookup/form/author.uri>

¹⁶ <https://isni.org>

PIDs as the strategy evolves. Finally, recording these comparisons and the criteria used helps to create an ‘organizational memory’ for the PID strategy initiative.

The matrix has been designed to match the internal characteristics of each PID to assessment criteria drawn from our consultations with Canadian research community stakeholders. It can be applied to newly emerging and well-established PIDs alike, and it focuses on a range of features of both the technology underpinning the PID service, and of the organizations providing those services. Since persistence is a feature of organizations or initiatives, rather than technology, the matrix also covers the governance, community engagement, and published policies of the parent organization.

The matrix provides a snapshot of the status of a PID at the time of its evaluation. It should, therefore, be treated as a living document, open to expansion or amendment as community needs or priorities evolve.

External factors, such as integration levels, adoption in specific communities, interoperability with other PIDs, policies relating to PIDs, etc., are not included in this matrix. These factors are highly changeable, for example, as the capabilities of platforms and PID registries continue to advance and as new policies emerge. They are therefore part of the context for strategic activities, rather than a feature of any given PID, so they should be investigated as part of the roadmap design process, for example, by evaluating the adoption levels of particular platforms or systems and their PID integrations.

The criteria used in the matrix build on existing work from previous activities in this project and in the field of PIDs more widely. They were guided by the vision, mission, and principles set out in section 3.1 above, and the evidence and inputs that shaped specific criteria are described in section 2, above. The nine areas of assessment, and the criteria to enable them to be assessed and compared, are as follows:

- **Coverage:** the entity or entities identified by the PID system, geographical and disciplinary focus
- **Openness:** the licensing and accessibility of the metadata, source code, and documentation of the PID service, plus the terms of access to any Application Programming Interface¹⁷ (API) and core PID functionality
- **Interoperability:** the use of recognized web standards, whether the PID itself is standardized, and the range of ways that metadata and/or objects identified can be accessed (such as via resolution¹⁸ or content negotiation¹⁹)
- **Descriptiveness:** the availability of a formal metadata schema, with the capacity to be updated, to show the provenance of claims made in that metadata, and to support relationships and/or version management
- **Governance:** a range of factors including published policies and principles, stakeholder representation, transparency, and mechanisms such as elections for board membership
- **Community:** the specific communities served, approaches to membership and participation, the structures and support available to enable community participation, and approaches to outreach and equity
- **Sustainability:** the income streams available to the PID service, not-for-profit or equivalent status, either established or planned sustainability models, and long-term stewardship, backup plans, etc.
- **Preservation:** the long-term plans available to describe how preservation will be managed and resourced and how responsibilities can be passed on in the event of organizational failure

¹⁷ <https://en.wikipedia.org/wiki/API>

¹⁸ <https://www.doi.org/the-identifier/resources/factsheets/doi-resolution-documentation>

¹⁹ <https://citation.crosscite.org/docs.html>

- **Support:** the scale, scope, and nature of the technical support available to members and end users, plus linguistic coverage

The criteria can be combined to assess more complex questions. For example, combining the answers under Support with the sections on Openness, Interoperability, and Descriptiveness would enable an assessment of the appropriateness of the support provided to the user and/or implementer communities.

3.2.5 PIDs for priority entities

The final decisions about which PIDs to prioritize should be made by the Canadian research community. However, given that the draft criteria for evaluation have been established, it makes sense to assess likely candidate PIDs for the priority entities identified during the course of this project.

Various PIDs are available and are at varying stages of maturity and adoption. Given their existing widespread adoption, some PIDs are clear contenders for certain entities. For other entities PID solutions are still being developed; these will need to be evaluated when they launch a service that Canada could use or deploy.

We have therefore clustered PIDs into three categories: 1) entities for which there are mature, widely adopted solutions available; 2) entities for which no clear candidate PIDs have emerged and additional investigations are needed; and 3) entities for which there is no clear option, at present and a ‘watching brief’ on developments is recommended.

In all cases, the focus should be on the use case(s) each PID is designed to serve, and how well that matches the current and foreseeable future needs of the Canadian research community. The PIDs discussed below are proposed for consideration and evaluation by the CPIDAC and others, but it is not an exhaustive list and will very likely be subject to change.

Priority entities for which mature candidate PIDs are available are:

- DMPs and similar records

The only widely adopted PID for DMPs and similar records at present is the Digital Object Identifier (DOI) provided by DataCite. Canadian researchers and institutions already have access to these DOIs through the national consortium, making this the best option available at present.

- Grants

PIDs for grants have not yet been widely adopted, although numerous funders use their own internal identifiers (which are not PIDs) for grants and similar funding awards. DOIs are available for grant awards, with established offerings from Crossref, and emerging offerings from DataCite. Since both use the same underlying PID standard, we recommend careful comparison of their use cases before recommending widespread adoption by funders and other research-supporting organizations in Canada. Pre-existing membership access to DataCite DOIs via the national consortium may also be a factor in the selection process.

- Outputs (data, articles, books, etc., as well as “non-traditional”)

DOIs are the most widely adopted PID for published outputs of many kinds, with a range of content types supported by both Crossref and DataCite. Specific disciplines may have a preference for a specific identifier system, often one which has been built to serve a particular database or collection. We therefore recommend evaluating the relevant DOI offering first.

Archival Resource Keys (ARKs) are in use for datasets, often earlier in the research lifecycle than, for example, for works in progress, so we recommend that they are evaluated alongside DOIs, to make sure that the use cases they serve can be understood and met.

- People

ORCIDs are far and away the most established PID for researchers and contributors to research. They are already well supported by the national consortium and, as the de facto standard, they are an obvious candidate for consideration.

Other PIDs for people include the International Standard Name Identifier (ISNI). This PID may serve wider use cases, especially for historical and cultural figures and for the library and archives community.

Proprietary identifiers, such as the Scopus Author ID are also widely adopted, but their use is limited to specific commercial systems and they do not, therefore, meet many of the criteria in the PID matrix.

Priority entities for which candidate PIDs are less clearcut, for a variety of reasons, are:

- Organizations (institutions, funders, etc.)

Adoption of Research Organization Registry (ROR) identifiers continues to increase in the research space. ROR's primary use case is to match individuals to their organizational affiliations, meaning that it will meet many of the needs identified in our community consultations. However, while ROR already provides good coverage of institutions, it does not yet fully cover funding organizations (pending the transition of data from Crossref's Open Funder Registry, expected to complete in 2025).

Two other widely available PIDs for organizations are ISNI (open) and the proprietary Ringgold identifier, which is mostly used by publishers to identify institutional subscribers.

- Facilities, instruments, equipment

These entities represent a very diverse and dispersed landscape of practices and expectations. Facilities may be standalone organizations (in which case we recommend using the same PID as for organizations, above), or hosted within larger organizations.

Instruments are currently served by one generalist PID, DataCite DOIs.

Research Resource Identifiers (RRIDs) are used for biological resources in some communities, but at present there is no widely adopted PID for equipment per se.

More work needs to be done to ensure that PID coverage is robust for these entities. Since a PID option exists for instruments, this should focus on adoption. A unified approach may not be possible for facilities and equipment but, if community consensus can be achieved, the next step would be to identify and evaluate existing or emerging PIDs that match that use case.

Finally, the **priority entities for which a watching brief on PID developments is recommended** for now are:

- Projects

Projects are a critical yet poorly understood entity within the research landscape, which are often seen as synonymous with grants. An important difference between the two is that projects can be seen as something one **does**, while grants are something one **gets** — and, in fact, a project may be supported by one grant, many grants, or no grants at all.

Currently, the Research Activity Identifier (RAiD) is the only viable candidate PID for projects. It is still at a relatively early stage of development, but a RAiD service may be possible in Canada within the next 12 months, so this should be a high priority watching brief. Rolling out the service and assuring rapid adoption and integration across the Canadian landscape will take a dedicated program of activity, so preparation for this is recommended.

- Software

There is not yet a clear PID option for software which, like facilities and equipment, is diverse and dispersed, with no standard practice at present. Software Hash Identifiers (SWHIDs)²⁰ show promise and with adoption and a more formalised metadata schema they have the potential to meet the needs articulated by Canadian stakeholders during this project. However, given their current level of maturity and adoption, we do not yet recommend a candidate PID for software, but recommend further investigation and an ongoing watching brief for developments in the field.

3.3 Landscape analysis

The success of the PID roadmap will depend on carefully targeted interventions, which reflect community priorities and serve practical and policy needs. To evaluate the current state of knowledge around such things as levels of PID adoption and awareness, willingness to invest time and resources in PID integration and outreach, or alignment with national objectives, we reviewed existing work from previous phases of the project, consortial outreach and surveys, and CPIDAC discussions and pooled this information to conduct a gap analysis. The findings of this analysis were then validated via community interviews and focus groups with funder and institutional representatives.

This information helps us to both identify actions to be taken in the roadmap, and to prioritize those with the greatest potential impact on the specific challenges facing the Canadian research community today.

3.3.1 Gap analysis

We combined the findings from previous phases of consultations in which we identified priority PIDs, and explored Canada's policy priorities through a mix of desk research and new findings from interviews, workshops, and focus groups to get a complete picture of PID activity and progress towards adoption. Beginning with cataloguing the already known technical, social and policy assets, we contextualized the Canadian research communities' top priorities based on our workshops. We identified components of the national PID strategy that are already in progress or being planned, and mapped them to priorities, needs, and dependencies that are not yet in place.

We looked at the policy, social, and technical components in the context of four themes:

1. Governance
 - a. Technical
 - b. Social
 - c. Policy
2. Technical capacity
3. Leadership and clarity
 - a. Strategy
 - b. Policy
 - c. Framework
 - d. Priorities

²⁰ <https://www.swhid.org/>

4. Community engagement and outreach

A strong, and well thought through governance structure is required due to the distributed nature of Canada's national PID infrastructure and strategy. Care must be taken to ensure that diverse communities are well represented and their needs adequately considered. At the same time, the need for diverse requirements to be considered must be balanced against the need for adequate standardization to ensure interoperability. Increasing the strength of governance will require the development of a more formal structure and will need to be adequately resourced.

Work on an overarching strategy for research management that builds on the vision, mission and principles is required. Good work has already been done towards realising this, and should continue. To underline that need for greater resources to be directed towards governance, a number of key decisions need to be taken with relative urgency. For example, does Canada want a national research management system similar to those used in Portugal or Czechia? Which PIDs are most important for researchers, institutions, and funders in Canada? What do the various communities in Canada, for example the francophone or indigenous communities, need from research infrastructures?

The policy work already undertaken to formulate the problem statement, PID pitch, and briefing document should be made use of to reach out to important stakeholders, decision makers, and policy makers in Canada. This is necessary to ensure diverse community engagement, raise awareness with decision makers and leaders, and to unlock funding. A strategic outreach campaign and engagement plan is required that covers both grass-roots awareness and leadership buy-in.

At the community level, the work of the ORCID and DataCite consortia has been very beneficial. There remain, however, gaps in knowledge, understanding and capability within the Canadian research management community, particularly with regard to the importance and value of PIDs, so efforts to build on the success of the two consortia need to continue. For instance, current research information systems (CRISs) are pieces of software that underpin much of the work of research managers, drawing resources together within universities and integrating with PIDs and their associated metadata registries. The creation of CRIS user groups will encourage communities of practice and help crystalize Canadian research management requirements. This will in turn empower the Canadian research management community, enabling them, for example, to encourage software developers and vendors to develop relevant and needed features.

Funders must lead by example and issue grant IDs as well as provide more visible political leadership in areas like technical and social change. Funders should also integrate ORCID and ROR into their funding systems. They have a high level of influence in all research ecosystems due to their ability to attach requirements to funding and to sponsor innovation, so the adoption of Grant IDs and integration of other PIDs, coupled with visible engagement and leadership would send a powerful message to the community.

Development of the technical aspects of the Canadian PID strategy will be very important for long-term success. Once a strategy has been developed and strong governance is put in place, it is important that technical capacity and capability are addressed so that institutions, funders, and other stakeholders are able to effectively support the development of adoption of PID services.

A centralized technical support service is needed to assist institutional PID adoption as well as reduce its cost and burden to individual institutions. Research has shown^{21,22} that central technical support

²¹ J. Brown, P. Jones, A. Meadows, and F. Murphy, 'Revised cost-benefit analysis for the UK PID Support Network', Zenodo, Nov. 2022. doi: 10.5281/ZENODO.7356219 .

²² Jones, P., Meadows, A., Murphy, F., Brown, J., Conway, A., & Griffith, L. (2024). Efficiency and insight: a cost-benefit analysis for a central service to support persistent identifier implementation in Ireland. Digital Repository of Ireland. <https://doi.org/10.7486/DRI.nz80kt123>

services to assist institutions in the technical processes, education, and change management of PID adoption reduces the overall system-wide cost of adoption. It is also likely to increase the pace of adoption and foster communities of practice leading to improved interoperability and information sharing.

Repository and CRIS plugins, APIs and other integrations should be supported with central funding once needs have been identified. There are technical challenges to be overcome in order for PID adoption to be highly effective. Since many institutions use similar software systems, there are economies of scale to be made by providing central funding to develop the necessary software extensions for integrations. Centrally developed plugins and integrations also improve levels of interoperability and standards alignment.

3.3.2 Institutions

We conducted two focus groups with representatives from Canadian research-performing organizations, on April 16 and 23 2024. With nine participants across the two meetings, attendance was lower than we had hoped, however, we gained some valuable insights, including, as one participant observed, the importance of looking at the opportunity cost — what will happen if there isn't a national strategy? Other key themes from the two meetings included:

Overview

The participants work in a variety of roles, including research data management, digital archiving, library technology, metadata/standards, and analysis. All were reasonably — in some cases very — familiar with PIDs. Most of them felt that the technology was not an issue but buy-in at both the leadership and the user level is a challenge; while some individuals are very enthusiastic, others are totally uninterested. This manifests itself in a lack of ownership within their institutions, making change challenging. Buy-in from underserved groups is also an issue; participants noted that the more the landscape is developed, the less likely it is to serve groups that have historically been underserved (or not served at all). This is both because developments often serve vested interests, and existing networks and service users will generally have the 'loudest voice'.

Leadership

Most participants were not aware of the CPIDAC, but they all saw a strong need for its role in terms of central coordination, information-sharing, corralling, etc. suggesting that raising awareness of the group and its role is clearly needed. Participants also noted that leadership is important in terms of ensuring that investment in PIDs is continuous and sustainable, i.e., persistent, which is again an opportunity for the CPIDAC. They also saw opportunities for key individual organizations to play more of a leadership role. For example, Tri-Agencies support would be a huge incentive for getting both researchers and institutions on board with adopting and implementing PIDs. As one participant asked, "If not the Tri-Agencies, then who?"

Community outreach and support

We need to ensure that the strategy is inclusive of, for example, non-traditional outputs, non-English speakers, Indigenous knowledge needs, etc., and that there is a clear value proposition for each stakeholder group. Central support and guidance are also seen as essential, to ensure consistency in the use and implementation of PIDs across the whole research community in Canada. This includes training and upskilling researchers, as well as accessible documentation — all in both English and French.

Privacy and security concerns

Concerns around privacy and security were raised by participants in both meetings, especially in relation to RAiD and the sharing of information at project level. "Any system that interoperates with other systems would need to be verified as technically secure".

Data quality and reliability

Concerns were also expressed around the trustworthiness of data; examples given included the reliability of information in an ORCID record, and link rot was also mentioned.

Key PIDs

There was some divergence in terms of which PIDs the participants would prioritize but ORCID was a clear favourite and is already the most highly used. PIDs for DMPs were also seen as a high or medium priority by many, but views on RAID were more mixed, in part because several participants were unfamiliar with it. Participants expressed a preference for community-driven PIDs, such as ROR. They mentioned some identifiers that are missing from the current priorities, including: ARKs (raised in relation to the commitment to persistence and a consequent emphasis on (meta)data quality); NOIDs (Nice Opaque IDentifier), a minter and name resolver for digitized content; and identifiers for subjects/subject codes.

Systems and workflows

Participants noted that many of their organizations do not have CRISs. They also felt that research evaluation workflows must be considered as an important part of the strategy.

3.3.3 Funders and agencies

Scheduling time with funders has been challenging in the past, and they are also often concerned about over-sharing information across agencies. We therefore gathered input and feedback from this group through six interviews (including one with two people from the same organization), conducted in mid to late April. Collectively, they represented federal and provincial, disciplinary-specific and generalist funders within Canada.

Overview

Most interviewees were not very familiar with the Canadian PID strategy and roadmap project, although they all had a fairly good understanding of PIDs in general. They engaged positively during the discussions and were willing to envisage a future in which funders' systems would be PID-enabled.

There were clear indications of frustration across the sector that PID adoption at their organizations has been lower and slower than the interviewees would have liked. They identified a number of sizable challenges that will need to be addressed in order for meaningful progress to be made in the medium term.

Current understanding and usage of PIDs

PIDs usage is sub-optimal in many of these agencies — while identifiers may be collected, they are not yet being used to connect and exchange information across systems. Internal conversations about ORCIDs are the most advanced, however, several interviewees reported that ORCIDs can only be added manually to their systems, with no further interoperability to enable the pushing or pulling of information to and from ORCID records. Internal identifiers (not PIDs) are used for grant application, tracking, and reporting purposes. FRQ is an exception: they have recently implemented ORCID single sign-on, are now working on pulling information from ORCID records, and will then start pushing information to them.

Interviewees engaged with — and easily understood the value proposition for — most of the PIDs mentioned in the interviews, and they expressed a particular interest in PIDs for data management plans, and RAIDs (although more clarity is needed about the difference between project and grant IDs).

PID awareness among the technical and knowledge-management focused interviewees was much higher than that of the programme managers and senior administrators. The technical staff see the scale of administrative burden, and understand how introducing PID-enabled workflows in grant

management platforms would be transformative. Their programme managers and administrators, on the other hand, are very aware of governmental requirements for handling data privacy and residency, and conflicts of interest; they are sensitive to the potential community pressures of imposing mandated use of PIDs on grantees. This results in a change-averse culture, with concerns about using organizational resources to invest in new systems.

Internal IDs for applicants/grantees are very common, and their use predates ORCID, so agencies are reluctant to move to ORCIDs (or PIDs generally). It is difficult for technical and operational colleagues to make the case for their organizations to integrate PIDs because their programme-facing, often more senior, colleagues don't understand the PID value proposition and may be reluctant to engage because of their concerns about compliance: introducing PIDs to these organizations will involve a transition period.

The funder landscape

All the interviewees recognized that resourcing issues in their organizations will need to be addressed in order for progress to be made. Getting technical development work prioritized is challenging, and there is also uncertainty about future governmental priorities, and whether their organizations can rely on the long-term support for PIDs needed for the level of investment required to be appropriate.

Several interviewees said that the federal government's Roadmap for Open Science²³ has helped with making the internal case for PIDs, however, there is clearly a tension between the national commitment to open science and the high-priority requirements for data residency, privacy and national security, licensing, and compliance. Funders are also very wary of accepting direct deposits of data, and about taking action towards increasing openness that may inadvertently jeopardize their ability to maintain trust in their other responsibilities — especially given that PID governance is outside of their direct control.

Grant management systems

Grant management systems across the funder landscape are a mess. For example, researchers who work across the Tri-Agencies have different accounts and identification numbers for each one, none of which are linked (the Tri-Agencies are, however, trying to commission a new system). Concerns were also raised about getting the grant system developers who serve funders to integrate PIDs.

Potential opportunities for action and interventions

Interviewees were very helpful in identifying opportunities to build awareness, adoption, practice, and trust among funders for PIDs. These are grouped thematically below.

Documented advice and guidance

Suggestions for materials which could help make the case, build internal buy-in, or allay current concerns included:

- Clear statements on the importance of PIDs from the funder's perspective
- Advice on implementations, integrations, best practices for systems, how to future-proof, etc.
 - Particularly on global alignment with other funders
 - Guidance on how to work with vendor user groups
- Explicit guidance on ensuring equitable engagement with Francophone communities
- Guidance on what to prioritize, especially since there is so much to do
- Resources to manage the internal selling
- Government/agency success stories, such as the Australian cost benefit analysis materials
- Evidence of how much easier it would be to get information if PIDs were implemented (e.g. with RAIDs)

²³ <https://science.gc.ca/site/science/en/office-chief-science-advisor/open-science/roadmap-open-science>

- The importance of identifying and supporting PID champions

Government and policy issues

These may require some additional consultation and/or engagement with the specific legislative landscape and funder priorities:

- Data has to be resident in Canada, so ORCID cannot be mandated. Requirements on security (in response to a recent case of Chinese scientists stealing data from the government) have been tightened. This has complicated the decision-making environment, but it could provide opportunities to build engagement among senior leaders within the funding agencies
- Work out how to support both narrative CVs and the CCV through PID-enabled workflows
- Build confidence that there is an exit strategy if needed
- Demonstrate investment rather than cost (the cost in this instance would be: what will happen if they do not take action)

Collaborations

Some possible convening opportunities:

- Work with librarians and organizations such as Érudit and Acfas on increasing understanding and uptake and, crucially, ensuring accessibility via non-English documentation
- Work with the DataCite and ORCID consortia to accelerate messaging via existing networks (several interviewees represented funders that are members of ORCID's Funder Interest Group which has committed to improving coordination around PIDs²⁴)

3.3.4 Key policy points

The funder interviewees drew clear connections between the Canadian federal government's direction of travel and their agencies' own internal views of their room for manoeuvre. The funders themselves also have some high-level strategic priorities, including changes to the Canadian Common CV (CCV). Further exploration of the touchpoints they identified (below) would help understand whether they are potential drivers or barriers to progress, as well as what the various stakeholders will need in order to move forward.

PIDs and existing funder priorities

All the funders we interviewed are pursuing specific objectives that could be woven into the PID narrative, including reducing burden; facilitating accessibility; and supporting multilingualism.

The CCV is viewed by researchers as a necessary evil — with a poor user interface, lack of interoperability across agencies, and high administrative burden. However, since the CCV will eventually be deprecated by the Tri-Agency, updates and interoperability may finally be possible. And other funders will be deciding on its alternative, which again provides an opportunity for a more PID-enabled, less burdensome replacement.

PIDs and existing government priorities

The interviewees explained that a clear government directive on what to do about PIDs would be helpful in unlocking and directing resources, as well as building confidence and trust in PID systems. An official response or manual which explains how PIDs intersect with the need for clarity on licensing, managing the various privacy requirements, and ensuring compliance around conflicts of interest and data security, would have considerable impact. It would be even more useful to highlight how PIDs can connect Canadian research and researchers with the international community.

²⁴ <https://info.orcid.org/wp-content/uploads/2024/02/ORCIDFigRecommendations2023.pdf>

The Canadian government is starting to roll out data reference standards (as of December 2023, beginning with a standard for Canadian provinces and territories). This is likely to include organizations, providing an opportunity to map these government organizational identifiers to ROR. Further investigation of these standards could help ensure that, wherever possible, PIDs are used to minimize the work and complexity, for example, by keeping the number of crosswalks between systems as low as possible.

Several interviewees noted that the Canadian government's open science roadmap²⁵ is potentially helpful. To achieve the overall objective of open science, and taking into consideration feedback from intradepartmental consultation (Recommendation 2), departments and agencies were tasked with delivering action plans²⁶ to show how they are planning to achieve the open science roadmap's recommendations for open science by June 2021. Not all the funders have published such plans, but NRC²⁷ has done so.

4. Recommendations

The recommendations set out here are categorized in line with the structure of the gap analysis. These categories are useful here in that they point to the kinds of expertise and project activities necessary to deliver each recommendation and could be used to structure programme design for the delivery of the roadmap.

4.1 Leadership and clarity

There is a need to establish strong and visible leadership. The distributed nature of PID infrastructure poses risks to momentum that can only be overcome through proactive engagement and coordination. While CPIDAC has done and is doing excellent work in keeping the PID agenda alive, it currently has an explicitly advisory role and no formal leadership remit.

PID adoption requires many separate groups to act in harmony, and since the benefits of PIDs are generated in large part by network effects, each group requires others to act if they are to receive the desired value from their investments in PID integration and adoption. This is a classic collective action problem, and will require a strong central political drive, ideally at the level of the Chief Scientific Advisor or federal governmental level to push this program higher in the list of competing organizational priorities. Without this level of 'ownership' of the strategy, it will not be possible to unlock the kinds of investment needed for consistent, equitable access to PIDs, cross-community coordination will be much harder to achieve, and ultimately there will be little to push PID integrations to the top of the 'to do' list, given the many competing priorities across the sector.

4.2 Governance

A roadmap to develop strong technical, social, and policy governance is urgently required. There are a number of key decisions that need to be made about how Canada wishes to handle research management information and there is an urgent need to consult traditionally underserved communities in Canada, such as Indigenous peoples and Francophone communities.

Building on structural, technical, and governance decisions, there is a need to prioritize which PIDs should be most strongly supported, first based on community consultations and the advice of a

²⁵ <https://science.gc.ca/site/science/en/office-chief-science-advisor/open-science/roadmap-open-science>

²⁶ <https://science.gc.ca/site/science/en/open-science-helping-make-science-accessible-all-canadians/departamental-open-science-action-plans>

²⁷ <https://nrc.canada.ca/en/corporate/planning-reporting/open-science-action-plan-response-government-canada-roadmap-open-science>

technical advisory committee. This prioritization should be considered a living document and constantly revised as the strategic research management needs of Canada evolve.

4.3 Technical capacity

There are a number of technical projects that could assist in driving PID adoption and usefulness. Many of these potential programmes, for example, plugins for repositories for specific PIDs, are dependent on technical and social governance decisions as described above. In the meantime, a central technical support service for PID adoption should be set up.

The community discussions highlight several PID integrations to prioritize. We have broken these out into three general categories, as they help to illustrate which stakeholder and system is indicated by each.

Inputs to research

Canadian funders should lead by example and begin issuing grant PIDs (in contrast to existing internal grant numbers) as soon as possible. This functionality would need to be incorporated into grant and award management systems. Further adoption in reporting systems and internal research databases would be highly recommended to ensure that the value of these PIDs can be delivered to funders as a return on their investment, and to maximize the reduction in administrative burden for researchers.

Research activities

Canada should adopt Research Activity Identifiers (RAIDs) for projects as soon as is feasible, with the caveat that the system is at an advanced stage of development and at the time of writing is scheduled for global roll-out in the second half of 2024. Meanwhile, a subgroup of the CPIDAC should investigate the options and requirements for a RAID Registration Agency for Canada or evaluate a partnership with one of the other potential North American Registration Agencies. An additional benefit of these PIDs is the transparency they could bring to global activities and collaborations.

Expand the funding and capacity of the ORCID-CA Consortium, possibly via additional staffing via the PID support service mentioned above, with a remit to improve ORCID integrations and consistency, and focus on populating ORCID records with trustworthy data. This implies integrations in institutional and funder systems in the first instance, with the capability to write to ORCID records as well as to ingest data from them.

Common, global, and open identifiers for organizations are needed, and should be strongly encouraged in research management processes, and potentially mandated for reporting. RORs are currently most used and should be evaluated for use for research organizations including funders, and with an eye to their potential use for identifying research facilities.

PIDs, such as DOIs, for instruments are urgently needed to track and manage investments, and for research integrity assurance. This will require integrations into institutional systems, both for the conduct of research, including e-lab notebooks, and for research management systems.

Research outputs

Throughout the landscape analysis, PIDs for datasets, primarily ARKs for works in progress and DOIs for static or published materials, emerged from discussions as an institutional and practitioner priority. As with instruments, this implies integrations in institutional systems, but with additional value to be derived from funder integrations in reporting and evaluation systems. This will also help to monitor progress towards policy goals around data availability, sovereignty, and re-use.

The use of DOIs for published outputs from all Canadian scholarly publishers should be strongly encouraged. Noting that smaller, emerging, or non-traditional publication venues may not be

resourced or equipped to access or leverage such PIDs, the CPIDAC should investigate the needs of smaller publishers and the possibility of a sponsorship or similar arrangement to improve access and coverage.

A common thread across all these actions is the need for institutions and funders to prioritize re-use of metadata from PID registries wherever possible to minimize duplication of effort and errors from manual data entry.

4.4 Community engagement and outreach

There is a need to get PIDs and research infrastructure higher up the national political agenda. Several deliverables, from this and previous work, including the problem statement, PID Pitch and briefing document, can form central components of messaging. A strategic outreach strategy is needed to make best use of these resources. To ensure Francophone engagement, a bespoke strategy may include work with additional stakeholders, such as Érudit and Acfas. There may be additional openly available MoreBrains outputs relating to PIDs not directly linked to this project which could usefully be translated into French.

In preparing outreach content, ensure that materials explain where the value lies in integrating PIDs. As an example, emphasizing that value is delivered from ORCID IDs by increasing the number of integrations and amount of time that is saved by reducing the administrative burden, rather than by how many ORCID IDs are registered.

Work is needed to directly address the situation for funders who are currently uncertain what the effects of introducing PIDs will be on their overarching responsibilities for data security, privacy, residency, etc. Where appropriate, include support from the Open Science Roadmap initiative. Funders are currently caught in stasis between these two imperatives, so insights into how PIDs could help with managing them could be transformative.

More community support is needed. The ORCID and DataCite consortia have done an excellent job of coordinating across institutions. Extending their capacity with additional support, and ensuring visibility for community activities with cross-stakeholder coordination, via for instance CRIS and repository user groups, national professional associations, will help to maintain momentum and propagate good practice.

Motivating and sustaining action will require evidence of the potential return on investments, and monitoring of the actual real-world delivery of those returns. Existing analyses, such as the MoreBrains Cost-Benefit analyses conducted in the UK²⁸, Australia²⁹, and Ireland³⁰, could be repurposed for Canadian audiences, and offer useful methodological approaches for assessing the scale and rate of adoption.

²⁸ Brown, J., Jones, P., Meadows, A., & Murphy, F. (2022). Revised cost-benefit analysis for the UK PID Support Network (Version 1). Zenodo. <https://doi.org/10.5281/zenodo.7356219>

²⁹ 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>

³⁰ Jones, P., Meadows, A., Murphy, F., Brown, J., Conway, A., & Griffith, L. (2024). Efficiency and insight: a cost-benefit analysis for a central service to support persistent identifier implementation in Ireland. Digital Repository of Ireland. <https://doi.org/10.7486/DRI.nz80kt123>

Investir dans un écosystème de recherche efficace et transparent au Canada

Données probantes, analyse et
recommandations pour
une feuille de route
des identifiants pérennes canadiens



7 juin 2024

Remerciements

Ce rapport a été réalisé à la demande de l'Alliance de recherche numérique du Canada (l'Alliance) et du Réseau canadien de documentation pour la recherche (RCDR), grâce à un financement de l'Alliance et à de précieux conseils du Comité consultatif canadien sur les identifiants pérennes (CCCPID).



**Digital Research
Alliance of Canada**

**Alliance de recherche
numérique du Canada**



**Canadian Research
Knowledge Network**

**Réseau canadien
de documentation
pour la recherche**

Les auteurs souhaitent remercier les partenaires du projet, les membres du CCCPID et tous les membres de la communauté de recherche canadienne qui ont si gentiment offert de leur temps et fait part de leurs réflexions dans le cadre de ce projet.

Ce rapport et les documents qui lui sont associés sont fournis sous une licence Creative Commons Attribution 4.0 International (<https://creativecommons.org/licenses/by/4.0/deed.fr>).

Sommaire

Cette phase du projet de feuille de route des PID canadiens repose sur les deux phases précédentes visant à élaborer une vision, une mission et des principes pour la stratégie, un ensemble de PID et d'intégrations prioritaires et un ensemble de recommandations concrètes pour la mise en œuvre de la stratégie.

Les principes soulignent que la communauté a besoin d'une stratégie qui donne la priorité aux systèmes ouverts, qui soit inclusive et équitable, qui encourage la mobilisation et la collaboration de la collectivité, qui ait une portée internationale, qui apporte des changements efficaces et tangibles, et qui soutienne des données de qualité, fiables et vérifiables.

Les consultations auprès de la communauté ont permis d'identifier huit entités susceptibles d'être évaluées et priorisées au cours de la prochaine phase de l'élaboration de la stratégie nationale sur les PID. Voici ces priorités, présentées en ordre alphabétique : la gestion des données de recherche (GDR) ; les installations, les instruments et les équipements ; les logiciels ; les organisations ; les personnes ; les projets ; les résultats ; les subventions.

Le projet a permis de créer une matrice de sélection des PID, destinée à faciliter une comparaison des PID potentiels pour ces entités, dans le but de favoriser une prise de décision éclairée tant aujourd'hui qu'à l'avenir, et à partir de critères qui intègrent les meilleures pratiques à l'échelle mondiale de même que les priorités de la communauté canadienne, exprimées dans la vision, la mission et les principes.

Bien que ce soit la communauté de recherche canadienne qui décidera finalement des PID à adopter, nous avons évalué les candidats probables pour les entités prioritaires. Certains PID sont déjà bien établis et largement adoptés (DOI pour la GDR, les subventions et les résultats ; ORCID pour les personnes). Il existe cependant plus d'un PID potentiel pour les organisations. Il existe des DOI pour les instruments, mais ils ne sont pas encore très répandus. Il n'y a pas d'options claires de PID pour les installations, les équipements et les logiciels. Il existe un PID possible pour les projets, le Research Activity Identifier, qui est en cours de développement.

En nous basant sur les travaux réalisés au cours des phases précédentes du projet, sur la sensibilisation et les enquêtes du consortium, et sur des discussions du CCCPID, nous avons procédé à une analyse des lacunes, qui a été validée par des entretiens avec des membres de la communauté et des groupes de discussion. Cela nous a permis d'identifier les éléments du cadre technique, social et politique qui doivent être harmonisés pour que la stratégie puisse être mise en œuvre. Quatre thèmes sont ressortis de l'analyse des lacunes : la gouvernance, la capacité technique, le leadership et la clarté, ainsi que la mobilisation et la sensibilisation de la communauté. Nos recommandations sont regroupées autour de ces thèmes.

Un leadership politique clair et de haut niveau s'impose pour mettre en œuvre la stratégie sur les PID, pour aider à dégager des ressources et pour placer l'intégration des PID au premier rang des priorités organisationnelles. La supervision de l'élaboration de la stratégie, notamment sur le plan de l'inclusion et de la représentation des collectivités et de la sélection des PID sur lesquels mettre l'accent, sera déterminante pour obtenir l'adhésion de la communauté. Les capacités techniques pour l'intégration des PID sont limitées, et nous proposons, pour chaque secteur de la communauté, des recommandations détaillées concernant les premières étapes ou les priorités. Enfin, nous insistons sur la nécessité d'une sensibilisation ciblée et d'un message cohérent pour les communautés importantes et pour les différents groupes de praticiens dans le domaine.

Table des matières

1. Contexte	27
2. Synthèse des données probantes	27
3. Constatations et résultats	29
3.1 Vision, mission et principes	29
3.1.1 Vision	29
3.1.2 Mission	30
3.1.3 Principes	30
3.2 Identification des entités prioritaires et des PID potentiels	31
3.2.1 Résultats de l'atelier de priorisation	32
3.2.2 Entités prioritaires	33
3.2.3 Sensibilisation aux PID et disponibilité	34
3.2.4 La matrice de sélection des PID	34
3.2.5 Des PID pour les entités prioritaires	36
3.3 Analyse contextuelle	38
3.3.1 Analyse des lacunes	39
3.3.2 Établissements	41
3.3.3 Bailleurs de fonds et agences	42
3.3.4 Points importants des politiques	45
4. Recommandations	46
4.1 Leadership et clarté	46
4.2 Gouvernance	46
4.3 Capacités techniques	46
4.4 Mobilisation et sensibilisation de la communauté	48

1. Contexte

Depuis 2021, l'équipe de MoreBrains Cooperative collabore avec le Réseau canadien de documentation pour la recherche (RCDR)³¹ afin d'aider à établir les bases d'une stratégie nationale sur les identifiants pérennes (PID) au Canada. Ce travail est financé par l'Alliance de recherche numérique du Canada (« l'Alliance »)³², et en est maintenant à sa troisième phase. La première phase a été largement consultative, et a permis de prendre en compte les travaux, les politiques, les collaborations et les consortiums existants afin de déterminer les mesures à prendre pour créer une stratégie complète. À partir de nos conclusions, nous avons, au cours de la deuxième phase, élaboré un ensemble de produits de communication, notamment des notes d'information pour les principales parties prenantes, des documents d'introduction pour les nouveaux venus dans le domaine des PID et des analyses des besoins et des défis particuliers de la communauté en matière d'adoption et d'intégration des PID. La troisième phase a été consacrée à l'élaboration d'actions et de priorités concrètes et précises pour l'intégration des PID, sur la base des résultats des consultations et des recherches menées au cours des phases précédentes. Parmi les résultats attendus figure une liste hiérarchisée de PID, d'intégrations de systèmes et d'activités de renforcement des capacités.

Pour réaliser ce travail, nous avons adopté une approche en trois étapes, avec l'objectif de proposer un ensemble d'actions prioritaires claires, fondées sur des données probantes, axées sur des défis précis et répondant à une demande explicite de la part de chaque groupe de parties prenantes :

1. Travailler en étroite collaboration avec l'équipe de projet et le Comité consultatif canadien sur les identifiants pérennes (CCCPID) pour élaborer une vision, une mission et un ensemble de principes pour la stratégie, ainsi que pour réaliser une analyse des lacunes en vue d'une évaluation des connaissances actuelles et des capacités du secteur ;
2. Identifier les PID et les intégrations prioritaires, en utilisant l'ébauche de la « matrice de sélection des PID » de la deuxième phase de ce projet, ainsi que les résultats de l'analyse des lacunes, et en faisant appel à des recherches complémentaires sur le contexte ;
3. Réaliser six entretiens avec des membres du personnel technique des organismes de financement et deux groupes de discussion avec des représentants institutionnels, conçus pour fournir des indications précises et exploitables dans le cadre d'un programme d'activité continu visant à concrétiser la vision stratégique.

2. Synthèse des données probantes

Les trois phases du projet nous ont permis d'accumuler un grand nombre de données probantes. Une grande partie du travail décrit dans ce rapport s'est concentrée sur le traitement de ces informations afin qu'elles puissent servir à la prise de décision. Il en résulte quatre éléments de synthèse distincts qui façonnent les recommandations formulées dans le présent rapport :

1. Une vision, une mission et des principes

La vision et la mission définissent les objectifs ultimes de la stratégie sur les PID, tandis que les principes orienteront la conception et la mise en œuvre de la stratégie, notamment pour la prise de décisions et l'établissement de priorités.

Ces déclarations s'appuient sur les entretiens, les groupes de discussion, les ateliers, les sondages et les autres activités de consultation de la communauté menés au cours des phases précédentes du projet et par les consortiums canadiens ORCID et DataCite. Les discussions avec le CCCPID sur les

³¹ <https://www.crkn-rcdr.ca/>

³² <https://alliancecan.ca/fr>

préoccupations et les aspirations de la communauté au cours des deux premières phases ont également été incluses dans l'analyse. Le réexamen des notes prises à l'époque et des résumés présentés dans nos rapports nous a permis de nous faire une idée précise des thèmes importants pour les membres de la communauté de recherche avec lesquels nous nous sommes entretenus. Et avant d'être finalisés, la vision, la mission et les principes ont fait l'objet de plusieurs cycles d'itérations en collaboration avec l'équipe du projet et le CCCPID. Pour plus de détails, veuillez consulter la partie 3.1 ci-dessous.

2. Un schéma du cycle de vie des informations sur la recherche

Une étape importante de ce processus consiste à comprendre les problèmes pratiques que les PID et les métadonnées qui leur sont associées pourraient résoudre dans un contexte de flux de travail de recherche. Au cours de la deuxième phase du projet, nous avons organisé deux ateliers virtuels qui ont abordé les questions relatives à l'information sur la recherche sous des angles différents, mais avec un point commun : la perte d'informations ou les dysfonctionnements dans l'échange de données.

Les résultats de ces ateliers ont ensuite été organisés sous la forme d'un schéma du cycle de vie de l'information sur la recherche, décomposant les activités qui se déroulent généralement à différentes étapes, depuis la conception des idées de recherche jusqu'à l'établissement de rapports, l'évaluation et l'étude d'impact après le projet.

Cette présentation des enjeux nous a permis de superposer les suggestions de la communauté concernant les étapes susceptibles d'être améliorées grâce à l'intégration des PID, et d'associer les interventions cohérentes ou à fort impact à la fois à la communauté et à l'activité qui en bénéficierait. Cela nous a également permis de définir avec précision les entités pour lesquelles des informations étaient nécessaires à chaque étape. Ces entités ont ensuite été considérées comme prioritaires pour une évaluation plus approfondie en matière de disponibilité et d'adéquation aux PID.

Les principales conclusions concernant les points névralgiques et les entités prioritaires sont exposées dans la partie 3.2.2.

3. Une matrice pour la sélection des PID

Pour faciliter l'évaluation des PID disponibles pour les entités prioritaires, nous avons besoin de critères cohérents et transparents, qui tiennent compte des valeurs et des objectifs énoncés dans la vision, la mission et les principes, et qui intègrent les priorités et les préoccupations de la communauté identifiées jusqu'à présent au cours du projet. Nous avons à nouveau analysé les résultats bruts des huit ateliers afin de définir les caractéristiques des PID qui reflètent les priorités de la communauté. Nous avons également réexaminé les synthèses précédentes de nos rapports de fin de phase ou des résultats particuliers du projet, comme le schéma du cycle de vie de l'information sur la recherche.

Les critères devaient également tenir compte des bonnes pratiques établies en matière de fourniture de PID, tant sur le plan technologique que sur celui de la gouvernance. Pour ce faire, nous nous sommes appuyés sur les travaux entrepris dans le cadre de projets antérieurs portant sur les PID, comme le projet ORCID and DataCite Interoperability Network (ODIN)³³, financé par la Commission européenne, la politique en matière de PID de l'European Open Science Cloud³⁴ et les Principles of Open Scholarly Infrastructure³⁵.

³³ <https://cordis.europa.eu/project/id/312788>

³⁴ <https://doi.org/10.2777/926037>

³⁵ <https://openscholarlyinfrastructure.org>

À partir de toutes ces données, nous avons élaboré 9 catégories, comprenant 39 critères distincts, dans le but d'éclairer la comparaison et la sélection des PID à inclure dans la stratégie. Les détails de ce processus sont présentés dans la partie 3.2.4, et une liste complète des données utilisées pour la création de la matrice est fournie dans le premier onglet de la feuille de calcul de la matrice de sélection des PID.

4. Une analyse des lacunes

La compréhension des atouts, des capacités et des systèmes déjà disponibles pour appuyer la mise en œuvre de la stratégie sur les PID, et leur mise en parallèle avec les lacunes connues ont également joué un rôle essentiel dans notre synthèse. Cela nous a aussi permis de dégager des éléments ou des mesures supplémentaires à prendre pour combler les lacunes, de mettre en évidence des dépendances et de formuler des recommandations concrètes en matière d'actions et d'interventions.

Toutes les activités et tous les résultats du projet à ce jour ont été intégrés dans l'analyse des lacunes. Cela nous a permis d'obtenir une vue d'ensemble complète et de haut niveau du contexte dans lequel la stratégie sur les PID sera mise en œuvre. L'analyse a été représentée graphiquement à l'aide de l'outil de tableau blanc virtuel Miro, et les résultats ont ensuite été classés et présentés sous la forme de diapositives. Les résultats sont décrits ci-dessous dans la partie 3.3.1.

3. Constatations et résultats

Ce rapport repose sur les résultats obtenus au cours des trois phases du projet réalisées à ce jour. Les consultations et les recherches menées au cours des deux premières phases ont été synthétisées sous la forme de résultats ciblés, chacun d'entre eux étant décrit dans la section correspondante ci-dessous. Le premier résultat de cette phase du projet a été l'énoncé sur la vision, la mission et les principes, qui sous-tend l'ensemble de la stratégie et qui, conjugué au flux de travail des entités prioritaires et à l'analyse de la deuxième phase, a façonné la conception de la matrice de sélection des PID et l'analyse des lacunes. Ces éléments ont à leur tour alimenté les questions de recherche que nous avons utilisées pour les entretiens et les groupes de discussion menés dans le cadre de l'analyse contextuelle.

Les recommandations découlant de ce travail ont été replacées dans le cadre de la vision, de la mission et des principes, afin de soutenir la prise de décision à venir et de garantir la clarté de l'orientation stratégique nécessaire à la réussite du projet.

3.1 Vision, mission et principes

La vision, la mission et les principes directeurs de la stratégie nationale canadienne sur les PID ont été établis sur la base des constats que nous avons faits lors des consultations de la communauté menées au cours des phases précédentes du projet, et élaborés en étroite consultation avec l'équipe de projet du RCDR et de l'Alliance. Une version préliminaire a ensuite été communiquée aux membres du CCCPID en prévision de leur réunion en personne à Ottawa, le 5 mars 2024 ; Alice Meadows, de MoreBrains, y a animé un atelier afin de recueillir d'autres commentaires et réactions de la part du groupe. Cet exercice a donné lieu à quelques modifications de la vision et de la mission, à la révision d'un des principes et à l'ajout d'un sixième principe. En voici les versions finales.

3.1.1 Vision

Une adoption généralisée des identifiants pérennes (PID) constituera le fondement d'un paysage efficace et équitable de données numériques de recherche au Canada, favorisant des liens ouverts et pérennes entre les intrants, les activités, les produits et les résultats de la recherche.

3.1.2 Mission

Sachant que les identifiants pérennes sont à la base d'un système de recherche numérique moderne, la stratégie nationale sur les PID fournira un cadre pour favoriser l'adoption généralisée des identifiants pérennes dans l'ensemble de l'écosystème de la recherche au Canada. Elle reflétera les priorités de toutes les communautés de recherche au Canada, en reconnaissant que la recherche s'effectue au sein et au travers de cultures, de disciplines et de langues différentes. Elle sera flexible, équitable et inclusive, et reposera sur des processus de consultation transparents. La stratégie sera fondée sur des données probantes – depuis l'analyse des besoins de la communauté jusqu'à l'évaluation de son efficacité – et s'adaptera et évoluera au fil du temps pour assurer sa viabilité sur les plans technique, organisationnel et social. Elle s'articulera autour d'un axe national, mais sera mise en réseau à l'échelle internationale, intégrant les meilleures pratiques où qu'elles se trouvent afin de soutenir l'utilisation des PID pour identifier de manière unique une série d'entités tout au long du cycle de vie de la recherche, favoriser l'interopérabilité entre les systèmes et les ressources d'information, et améliorer le flux de (méta)données précises.

3.1.3 Principes

Ces six principes ont été élaborés en réponse aux consultations menées auprès de la communauté en 2022 et 2023. Ils se veulent un guide pour l'élaboration et l'exécution de la stratégie nationale canadienne sur les PID, afin d'aider à orienter la prise de décision et à faire en sorte que la stratégie demeure axée sur les besoins et les priorités de la communauté au fil de son évolution future.

1^{er} principe – Les systèmes ouverts seront privilégiés dans la mesure du possible.

La valeur des PID réside en partie dans leur capacité à mettre en relation des informations contenues dans différents systèmes et entre différentes organisations. L'utilisation des PID, des métadonnées et des registres comme passerelles entre des systèmes d'information suppose un accès optimal à ces outils et le moins de limites possible à la réutilisation des données. Nous privilégierons donc les systèmes qui fournissent des interfaces de programmation d'applications (API) ouvertes, des licences claires et permissives pour l'accès aux données et leur utilisation, et qui disposent d'une gouvernance communautaire garantissant leur accès ouvert à l'avenir.

2^e principe – La stratégie, les activités qui en découlent et leurs résultats doivent être inclusifs et équitables.

Les avantages de la stratégie sur les PID doivent atteindre chacune des communautés engagées dans la recherche au Canada. Sachant que la valeur des PID augmente à mesure qu'ils sont utilisés, nous nous efforcerons de veiller à ce que la valeur générée par l'adoption des PID ne soit pas concentrée sur les groupes et les organisations qui bénéficient déjà d'avantages relatifs en matière de richesse et de statut. En pratique, cela signifie qu'il faut identifier les parties prenantes les moins favorisées et travailler activement pour obtenir des ressources, fournir un soutien ou des services supplémentaires et créer des trousseaux d'outils accessibles pour ceux et celles qui en ont besoin.

3^e principe – La mobilisation et la collaboration de la communauté seront un élément essentiel de la stratégie.

Notre stratégie est le fruit de consultations et de recherches auprès de la communauté. Nous veillerons à ce que cette approche axée sur la communauté continue d'évoluer, en encourageant de manière proactive la collaboration et les échanges entre les groupes de parties prenantes, y compris ceux qui ne font pas partie de la communauté actuelle des PID, afin de comprendre et de faire valoir leurs besoins. Nous ferons mieux connaître les PID et leurs avantages à tous ceux et celles qui contribuent à la recherche au Canada, grâce à des efforts de sensibilisation conçus pour mobiliser, expliquer et éduquer.

4^e principe – La stratégie nationale aura une portée internationale.

Nous reconnaissons que les activités de recherche canadiennes ne se déroulent pas uniquement à l'intérieur de nos frontières. Les chercheurs canadiens travaillent avec des collaborateurs du monde entier, reçoivent du financement d'organismes basés à l'étranger et publient dans des espaces gérés par d'autres pays. Nous représenterons activement notre pays au sein de ce réseau international de partenariats et nous accorderons la priorité aux PID, schémas et infrastructures utilisés à l'échelle internationale, afin de garantir que les systèmes nationaux et internationaux puissent fonctionner ensemble et échanger des informations sur les activités de recherche, peu importe où celles-ci se déroulent.

5^e principe – La stratégie vise un changement efficace et tangible.

L'adoption et l'intégration des PID dans les outils et les systèmes utilisés quotidiennement par les chercheurs et ceux qui les appuient nécessitent des investissements et des adaptations. Ces outils et ces systèmes doivent apporter une réelle valeur ajoutée, et permettre aux chercheurs d'être plus efficaces en améliorant la découverte, en facilitant leur vie professionnelle et en réduisant leurs tâches administratives. Concrètement, chaque intervention proposée dans la stratégie visera un défi ou une opportunité bien précise, avec des objectifs clairs et réalisables, et des critères validés par la communauté pour en évaluer l'impact.

6^e principe – La stratégie favorisera des données de qualité, fiables et vérifiables.

Des données et métadonnées de qualité et vérifiables permettent une analyse plus éclairée, améliorent la confiance grâce à leur fiabilité et rattachent les données à leurs créateurs, augmentant ainsi l'interopérabilité et l'intégrité des informations dans les différents systèmes. La stratégie sur les PID donnera la priorité aux systèmes qui favorisent la confiance et améliorent la qualité des PID et de leurs métadonnées, tout en respectant la sécurité et la protection de la vie privée des chercheurs et des contributeurs.

3.2 Identification des entités prioritaires et des PID potentiels

Le rapport *Towards a national PID strategy for Canada – Vers une stratégie nationale sur les PID pour le Canada*³⁶ présente les résultats de l'enquête menée par MoreBrains³⁷ sur le paysage actuel des PID au Canada et les attentes des communautés sur ce qu'une stratégie sur les PID doit leur apporter.

L'analyse exposée dans cette section répond en partie aux recommandations 3 et 4 :

3. Analyser les besoins et les obstacles identifiés par la collectivité des PID (p. ex., pour comprendre les raisons qui sous-tendent la non-participation) ;

4. Identifier un ensemble d'objectifs essentiels à atteindre pour la communauté et déterminer les PID qui y sont associés ainsi que les besoins en matière d'intégration.

En avril et mai 2023, en partenariat avec les coordonnateurs de projet du RCDR et de l'Alliance, MoreBrains a organisé deux ateliers virtuels avec des membres de la communauté. Le premier était axé sur l'échange d'informations entre les organismes de recherche (comme les universités et les instituts de recherche) et les bailleurs de fonds pour la recherche ; le second était consacré à l'intégration et à l'adoption des PID au sein des établissements.

³⁶ Josh Brown, Phil Jones, Alice Meadows et Fiona Murphy (2022). *Towards a national PID strategy for Canada - Vers une stratégie nationale sur les PID pour le Canada*. Zenodo. <https://doi.org/10.5281/zenodo.7217469>

³⁷ <https://www.morebrains.coop/>

3.2.1 Résultats de l'atelier de priorisation

Les conclusions du premier atelier sont présentées ci-dessous et au point 3.2.2. Elles décrivent respectivement les défis et les points névralgiques relevés par les participants, ainsi que les entités pour lesquelles une adoption complète de PID serait des plus utiles pour régler ces problèmes. Parmi les participants figuraient des représentants de bailleurs de fonds, dont le groupe des trois organismes et le Fonds de recherche du Québec, ainsi que des représentants de six universités.

La partie 3.2.3 présente les résultats du deuxième atelier, donne quelques orientations préliminaires pour des études supplémentaires et fournit des indications sur l'état actuel de l'adoption et de l'intégration des PID au Canada. L'atelier a réuni un large éventail de spécialistes des produits, des systèmes et de l'analyse, issus de bailleurs de fonds, de fournisseurs de services ou de plateformes et d'universités.

Défis et points névralgiques

Enjeux signalés du début à la demande de financement

Préparation de la demande de financement

- Assembler et structurer les données de CV
- Saisir des informations similaires dans plusieurs formulaires
- Difficultés liées à la réutilisation des informations dans les différentes demandes de subvention

Évaluation interne

- Approches ponctuelles ou non standard pour le partage des données relatives aux demandes
- Difficultés à identifier l'expertise appropriée et à éviter les conflits d'intérêts

Demande/soumission

- La saisie manuelle et répétitive des données engendre des erreurs et fait perdre du temps aux chercheurs
- Les demandes de subvention complexes comportent plusieurs situations d'échange d'informations avec un risque de perte de données
- Diversité et variations dans les données collectées
- Espace limité pour énumérer les résultats, les collaborations, etc.

Examen par les pairs

- Identification des conflits d'intérêts potentiels
- Les renseignements concernant le processus d'évaluation par les pairs ne sont pas visibles
- Dépendance excessive à l'égard d'un petit groupe de relecteurs

Enjeux signalés du financement à l'exécution du projet

Récompense et acceptation formelles

- Registres ambigus et inexacts des organisations et des personnes
- Numéros de subvention ou de bourse dont le caractère unique n'est pas garanti et auxquels aucune information descriptive n'est formellement associée
- Double saisie des données et dossiers discordants dans différentes organisations

Mise en place du projet

- Saisie manuelle des données avec peu d'interopérabilité ou de possibilité de réutilisation
- Difficulté à suivre les entités entre les systèmes ou à faire circuler l'information
- Difficultés liées au suivi de la participation et à la reconnaissance des contributions
- Approches ponctuelles ou non standard pour le partage des données de projet

Suivi de projet

- Il est difficile de suivre l'évolution dans le temps des résultats sans PID
- Remerciements incohérents et incomplets
- Renseignements exclusifs, incomplets ou hypothétiques
- Double saisie des données et dossiers discordants dans différentes organisations

Enjeux liés à l'établissement de rapports et à l'évaluation

Suivi des produits et résultats

- Difficulté à vérifier la conformité des résultats avec la politique
- Des conventions de dénomination et de description variables rendent les données difficiles à utiliser
- La collecte des données auprès des partenaires est un travail difficile et manuel
- Difficulté à comprendre les trajectoires professionnelles et les impacts

Rapports et conformité

- Rassembler des données qualitatives exige beaucoup de temps
- Reconnaissances incohérentes et incomplètes
- Difficile de reconstituer rétrospectivement les informations sur ce qui a été publié

Impact et évaluation

- Échange de données et interopérabilité médiocres entre les systèmes et les organisations
- Dépendance à l'égard d'agrégateurs/de bases de données tiers (souvent exclusifs et coûteux)
- Difficulté à comprendre les trajectoires professionnelles et les impacts
- Difficile de rattacher les impacts sociaux ou économiques aux activités/résultats du projet

3.2.2 Entités prioritaires

Si certaines entités peuvent être identifiées à partir des interventions relatives aux PID proposées par les participants à l'atelier, d'autres apparaissent implicitement dans les problèmes soulevés ou les solutions proposées. Par exemple, « réutiliser les données des registres de PID » a été suggéré comme solution à de nombreux exemples de difficultés liées à la collecte et à la réutilisation des données. Dans ces cas, nous avons déduit les PID les plus pertinents en analysant le contexte (comme la gestion de projet ou la demande de subvention).

Après avoir rassemblé toutes les entités référencées ou implicites dans le flux de travail en une seule liste dédoublée, les huit entités suivantes (présentées en ordre alphabétique) sont susceptibles d'être évaluées et classées par ordre de priorité lors de la prochaine étape de l'élaboration de la stratégie nationale sur les PID :

- Installations, instruments, équipements
- Logiciels
- Organisations (établissements, bailleurs de fonds, etc.)
- Personnes
- PGD et registres similaires
- Projets
- Résultats (données, articles, livres, etc., ainsi que les résultats « non traditionnels »)
- Subventions

Parallèlement à l'établissement des priorités de la communauté pour ces entités, nous devons analyser les options de PID disponibles et leur degré de maturité, et en tenir compte lorsque nous déciderons des principaux PID à privilégier dans le cadre de la stratégie sur les PID. Par exemple,

l'identifiant ORCID³⁸ (Open Researcher and Contributor Identifier) est une norme mondiale bien établie et incontournable pour l'identification des chercheurs, avec un consortium national et une communauté mobilisée au Canada. En revanche, les identificateurs de logiciels et de projets sont à un stade de développement beaucoup moins avancé sur le plan de l'infrastructure, de la gouvernance et de la communauté.

3.2.3 Sensibilisation aux PID et disponibilité

Le deuxième atelier a permis de faire la lumière sur les niveaux actuels de sensibilisation aux PID couramment utilisés dans les organismes de recherche, chez les bailleurs de fonds et auprès des fournisseurs de plateformes/systèmes.

Au sein des trois groupes, les identifiants ORCID pour les personnes, les identifiants ROR (Research Organization Registry)³⁹ pour les organisations et les DOI (identificateurs d'objet numérique, ou Digital Object Identifiers)⁴⁰ pour les résultats (en particulier les articles de revues et les ensembles de données) sont bien connus. Dans le groupe des fournisseurs de plateformes/systèmes, ces PID ont été les seuls à être discutés, à l'exception d'une seule mention du numéro ISSN (International Standard Serials Number)⁴¹ pour les revues. Les bailleurs de fonds ont également mentionné le Research Activity ID (RAiD)⁴² pour les projets, ainsi que des identifiants pour les instruments et les Research Resource IDs (RRID)⁴³. Les représentants des organismes de recherche ont également cité ces PID, ainsi que d'autres identifiants pour les organisations et les personnes comme l'identifiant exclusif Ringgold⁴⁴ ou l'identifiant d'auteur Scopus (Scopus Author ID)⁴⁵.

Il apparaît évident qu'au Canada, comme dans plusieurs autres pays, on connaît mal la disponibilité, la maturité et le statut des PID pour un grand nombre des entités énumérées au point 3.2.2. Il s'agit là d'un élément important, car il n'existe actuellement aucun PID adopté à grande échelle pour certaines entités (p. ex., les projets), alors que ces PID sont absolument essentiels pour résoudre bon nombre des problèmes de flux de travail soulevés au cours de l'atelier.

3.2.4 La matrice de sélection des PID

Pour plusieurs des entités prioritaires, il existe plus d'un PID potentiel, chacun présentant diverses caractéristiques qui les distinguent. Par exemple, les PID pour les chercheurs comprennent à la fois des PID exclusifs, comme l'Author ID de Scopus, et des PID ouverts, tels qu'ORCID ou ISNI⁴⁶. Parmi les PID libres, l'ISNI peut être utilisé pour tout ce qui porte un nom (y compris, par exemple, des personnages de fiction), tandis que l'identifiant ORCID est axé sur le secteur de la recherche.

Il est indispensable d'adopter une approche cohérente et coordonnée de l'adoption des PID à l'échelle nationale pour permettre à ceux-ci d'atteindre un niveau d'utilisation tel que le plus grand nombre possible de parties prenantes puissent bénéficier des avantages qu'ils procurent. La cohérence apporte également d'autres avantages, tels que la connaissance de l'ensemble des caractéristiques d'un PID donné, ce qui permet à la communauté de faire une utilisation optimale de ces PID dans les systèmes d'information sur la recherche.

³⁸ <https://orcid.org>

³⁹ <https://ror.org>

⁴⁰ <https://www.doi.org/>

⁴¹ <https://issn.org>

⁴² <https://raid.org>

⁴³ <https://scicrunch.org/resources>

⁴⁴ <https://www.ringgold.com/>

⁴⁵ <https://www.scopus.com/freelookup/form/author.uri>

⁴⁶ <https://isni.org>

Nous avons donc élaboré une matrice de sélection des PID afin de pouvoir les comparer facilement, sur la base d'une compréhension claire de leurs caractéristiques et de leurs spécificités. Cette matrice permettra de prendre des décisions éclairées sur la sélection d'un sous-ensemble parmi les PID potentiels disponibles, pour utiliser au mieux les ressources limitées dont nous disposons. L'adoption d'une approche structurée pour ces comparaisons permettra de les revoir, de les adapter ou de les mettre à jour ; elle facilitera également l'évaluation de nouveaux PID et/ou d'autres PID au fur et à mesure de l'évolution de la stratégie sur les PID. Enfin, consigner ces comparaisons et les critères utilisés contribuera à créer une « mémoire organisationnelle » pour le projet de stratégie sur les PID.

La matrice a été conçue pour faire correspondre les caractéristiques internes de chaque PID à des critères d'évaluation tirés de nos consultations avec les parties prenantes de la communauté de recherche canadienne. Elle peut s'appliquer aussi bien aux PID nouvellement créés qu'aux PID bien établis, et met l'accent sur une série de caractéristiques de la technologie qui sous-tend le service de PID, ainsi que des organisations qui fournissent ces services. La pérennité étant une caractéristique non pas des technologies mais des organisations ou des initiatives, la matrice couvre également la gouvernance, la mobilisation de la communauté et les politiques de publication de l'organisation mère.

La matrice donne un instantané de la situation d'un PID au moment de son évaluation. Elle doit donc être considérée comme un document dynamique, susceptible d'être complété ou modifié en fonction de l'évolution des besoins et des priorités de la communauté.

Cette matrice ne tient pas compte des facteurs externes, comme les niveaux d'intégration, l'adoption par des communautés données, l'interopérabilité avec d'autres PID, les politiques relatives aux PID, etc. Ces facteurs sont très changeants, par exemple, au fur et à mesure que les capacités des plateformes et des registres de PID progressent et que de nouvelles politiques apparaissent. Ils font donc partie du contexte des activités stratégiques, et ne sont pas une caractéristique d'un PID donné. Ils devraient donc être étudiés dans le cadre du processus de conception de la feuille de route, notamment en évaluant les niveaux d'adoption de plateformes ou de systèmes particuliers et leurs intégrations dans les PID.

Les critères utilisés dans la matrice s'appuient sur les travaux existants dans le cadre d'activités antérieures de ce projet et, plus largement, dans le domaine des PID. Ils ont été inspirés par la vision, la mission et les principes énoncés ci-dessus au point 3.1, et les éléments et informations qui ont façonné les critères particuliers sont décrits dans la partie 2 ci-dessus. Voici donc les neuf domaines d'évaluation, ainsi que les critères qui permettent de les évaluer et de les comparer :

- **Couverture** : l'entité ou les entités identifiées par le système de PID, la portée sur le plan géographique et disciplinaire
- **Ouverture** : les licences et l'accessibilité des métadonnées, du code source et de la documentation du service de PID, ainsi que les conditions d'accès à un interface de programmation d'application (API)⁴⁷ et aux fonctionnalités de base du PID
- **Interopérabilité** : l'utilisation de normes web reconnues, le fait que le PID lui-même soit normalisé, et l'éventail des modes d'accès aux métadonnées et/ou aux objets identifiés (p. ex., via la résolution⁴⁸ ou la négociation de contenu⁴⁹)
- **Caractère descriptif** : La disponibilité d'un schéma de métadonnées formel, avec la capacité d'être mis à jour, de montrer la provenance des affirmations contenues dans ces métadonnées, et de prendre en charge la gestion des relations et/ou des versions

⁴⁷ https://fr.wikipedia.org/wiki/Interface_de_programmation

⁴⁸ <https://www.doi.org/the-identifier/resources/factsheets/doi-resolution-documentation>

⁴⁹ <https://citation.crosscite.org/docs.html>

- **Gouvernance** : une série de facteurs comprenant les politiques et les principes publiés, la représentation des parties prenantes, la transparence et les mécanismes tels que les élections des membres du conseil d'administration
- **Communauté** : les communautés précises desservies, les approches en matière d'adhésion et de participation, les structures et le soutien disponibles pour permettre la participation de la communauté, et les approches en matière de sensibilisation et d'équité
- **Durabilité** : les flux de revenus dont dispose le service de PID, le statut d'organisme à but non lucratif ou l'équivalent, les modèles de durabilité établis ou prévus, et l'intendance à long terme, les plans de sauvegarde, etc.
- **Préservation** : les plans à long terme disponibles pour décrire les modalités de gestion et de financement de la préservation, et la manière dont les responsabilités peuvent être transmises en cas de défaillance de l'organisation.
- **Soutien** : l'ampleur, la portée et la nature du soutien technique offert aux membres et aux utilisateurs finaux, ainsi que la couverture linguistique

Les critères peuvent être combinés pour évaluer des questions plus complexes. Par exemple, en combinant les réponses relatives au soutien avec les sections sur l'ouverture, l'interopérabilité et le caractère descriptif, on pourrait évaluer l'adéquation du soutien fourni aux communautés qui utilisent et/ou mettent en œuvre le PID.

3.2.5 Des PID pour les entités prioritaires

Les décisions finales relatives aux PID à prioriser devraient être prises par la communauté de recherche canadienne. Toutefois, compte tenu du fait que les critères d'évaluation provisoires ont été établis, il est logique d'évaluer les PID potentiels pour les entités prioritaires identifiées au cours de ce projet.

Plusieurs PID sont disponibles et se trouvent à des stades divers de développement et d'adoption. Compte tenu de leur adoption généralisée actuelle, certains PID sont des solutions incontournables pour certaines entités. Pour d'autres entités, les options de PID sont encore en cours d'élaboration ; elles devront être évaluées lorsqu'elles lanceront un service susceptible d'être utilisé ou déployé par le Canada.

Nous avons donc regroupé les PID en trois catégories : 1) les entités pour lesquelles il existe des solutions éprouvées et adoptées à grande échelle ; 2) les entités pour lesquelles aucun PID potentiel ne s'est clairement démarqué et pour lesquelles des recherches supplémentaires s'avèrent nécessaires ; et 3) les entités pour lesquelles il n'existe pas d'option claire à l'heure actuelle et pour lesquelles il est recommandé de suivre de près leurs développements.

Dans tous les cas, l'accent doit être mis sur le ou les cas d'utilisation pour lesquels chaque PID est conçu, et sur leur adéquation avec les besoins actuels et futurs prévisibles de la communauté de recherche canadienne. Les PID présentés ci-dessous sont proposés à l'examen et à l'évaluation du CCCPID et d'autres organismes, mais il ne s'agit pas d'une liste exhaustive et elle sera fort probablement amenée à évoluer.

Entités prioritaires pour lesquelles des PID potentiels et éprouvés sont disponibles

- Les PGD et registres similaires

Le seul PID largement adopté à l'heure actuelle pour les PGD et les registres similaires est l'identificateur d'objet numérique (DOI) fourni par DataCite. Les chercheurs et établissements canadiens ont déjà accès à ces DOI par l'intermédiaire du consortium national, ce qui en fait la meilleure option disponible à l'heure actuelle.

- Les subventions

Les PID pour les subventions n'ont pas encore été adoptés à grande échelle, même si de nombreux bailleurs de fonds utilisent leurs propres identifiants internes (qui ne sont pas des PID) pour les subventions et autres bourses de financement semblables. Des DOI sont disponibles pour les subventions, avec des offres établies de Crossref et des offres émergentes de DataCite. Puisque les deux utilisent la même norme de PID sous-jacente, nous conseillons une comparaison minutieuse de leurs cas d'utilisation avant de recommander une adoption généralisée par les bailleurs de fonds et d'autres organismes de soutien à la recherche au Canada. L'accès préalable des membres aux DOI de DataCite par l'intermédiaire du consortium national pourrait également peser dans le processus de sélection.

- Les résultats (données, articles, livres, etc., ainsi que les résultats « non traditionnels »)

Les DOI sont les PID les plus largement adoptés pour les publications de toutes sortes, avec un éventail de types de contenu pris en charge à la fois par Crossref et par DataCite. Des disciplines particulières peuvent avoir une préférence pour un système d'identification donné, souvent conçu pour alimenter une base de données ou une collection particulière. Nous recommandons donc d'évaluer d'abord l'offre de DOI pertinente.

Les identifiants Archival Resource Keys (ARK) sont utilisés pour les ensembles de données, souvent plus tôt dans le cycle de vie de la recherche que, par exemple, pour les travaux en cours. Nous recommandons donc qu'ils soient évalués en même temps que les DOI, afin de s'assurer que les cas d'utilisation auxquels ils répondent puissent être compris et satisfaits.

- Les personnes

Les identifiants ORCID sont de loin les PID les plus reconnus pour les chercheurs et les personnes qui contribuent à la recherche. Ils sont déjà bien soutenus par le consortium national et, en tant que norme *de facto*, ils sont un candidat évident à prendre en considération.

Il existe d'autres PID pour désigner des personnes, notamment l'ISNI (International Standard Name Identifier). Ce PID peut servir à des fins plus larges, notamment pour les personnages historiques et culturels, et pour la communauté des bibliothèques et des archives.

Les identifiants exclusifs, comme l'identifiant d'auteur Scopus, sont également très répandus, mais leur utilisation est limitée à des systèmes commerciaux bien précis et ils ne répondent donc pas à la plupart des critères de la matrice des PID.

Entités prioritaires pour lesquelles les PID potentiels sont moins évidents, pour diverses raisons

- Les organisations (établissements, bailleurs de fonds, etc.)

L'adoption des identifiants du Research Organization Registry (ROR) continue de progresser dans le domaine de la recherche. Le ROR est principalement utilisé pour faire correspondre des personnes à leurs affiliations organisationnelles, ce qui signifie qu'il répondra à de nombreux besoins identifiés lors de nos consultations avec la communauté. Toutefois, si l'identifiant ROR couvre déjà bien les établissements, il ne couvre pas encore totalement les organismes de financement (en attente de la transition des données de l'Open Funder Registry de Crossref, qui devrait s'achever en 2025).

Il existe deux autres PID largement disponibles pour les organisations : l'ISNI (ouvert) et l'identifiant exclusif Ringgold, qui est principalement utilisé par les éditeurs pour identifier les abonnés institutionnels.

- Les installations, instruments, équipements

Ces entités représentent un environnement de pratiques et d'attentes très diversifié et fragmenté. Les installations peuvent être des organisations autonomes (dans ce cas, nous recommandons d'utiliser le même PID que pour les organisations, ci-dessus) ou être hébergées au sein d'organisations plus importantes.

Les instruments sont actuellement desservis par un PID généraliste, les DOI de DataCite.

Les identifiants RRID (Research Resource Identifiers) sont utilisés pour les ressources biologiques dans certaines communautés, mais il n'existe pas à l'heure actuelle de PID adopté à grande échelle pour les équipements à proprement parler.

Il faudra poursuivre les travaux pour s'assurer que la couverture des PID est solide pour ces entités. Puisqu'il existe une option de PID pour les instruments, l'accent devrait être mis sur son adoption. Une approche unifiée n'est peut-être pas possible pour les installations et les équipements mais, si un consensus pouvait être atteint au sein de la communauté, l'étape suivante consisterait à identifier et à évaluer les PID existants ou émergents qui correspondent à ce cas d'utilisation.

Enfin, les entités prioritaires pour lesquelles il est pour l'instant recommandé de suivre l'évolution des PID

- Les projets

Les projets sont une entité essentielle mais mal comprise dans le paysage de la recherche, et sont souvent considérés comme synonymes de subventions. Il existe une différence importante entre les deux : les projets peuvent être considérés comme quelque chose que l'on **fait**, alors que les subventions sont quelque chose que l'on **obtient**. En effet, un projet peut être soutenu par une subvention, par plusieurs subventions ou par aucune subvention.

Actuellement, le Research Activity Identifier (RAiD) est le seul PID potentiel viable pour les projets. Il en est encore à un stade de développement relativement précoce, mais un service de RAiD pourrait être possible au Canada dans les 12 prochains mois, ce qui devrait faire l'objet d'une surveillance hautement prioritaire. Le déploiement du service et la garantie d'une adoption et d'une intégration rapides dans l'ensemble du paysage canadien exigeront un programme d'activité approprié. Il est donc recommandé de s'y préparer.

- Les logiciels

Il n'existe pas encore d'option claire en matière de PID pour les logiciels. Comme les installations et les équipements, ceux-ci sont diversifiés et disparates, et ne font l'objet d'aucune pratique normalisée à l'heure actuelle. Les « Software Hash Identifiers » (SWHID)⁵⁰ et, avec une adoption plus large ainsi qu'un schéma de métadonnées davantage formalisé, ils pourraient répondre aux besoins exprimés par les parties prenantes canadiennes dans le cadre de ce projet. Toutefois, compte tenu de leur niveau actuel de développement et d'adoption, nous ne recommandons pas encore un PID potentiel pour les logiciels, mais nous conseillons de poursuivre les recherches et de rester attentifs à l'évolution de la situation dans ce domaine.

3.3 Analyse contextuelle

Pour que cette feuille de route des PID soit couronnée de succès, il convient de mener des interventions soigneusement ciblées, qui reflètent les priorités de la communauté et répondent à des besoins pratiques et politiques. Pour évaluer l'état actuel des connaissances sur des sujets tels que

⁵⁰ <https://www.swhid.org/>

les niveaux d'adoption des PID et de sensibilisation à leur sujet, la volonté d'investir du temps et des ressources dans l'intégration et la diffusion des PID, ou encore l'alignement sur les objectifs nationaux, nous avons passé en revue les travaux réalisés lors des phases précédentes du projet, les actions de diffusion et les enquêtes des consortiums, ainsi que les discussions du CCCPID. Nous avons ensuite regroupé ces informations pour réaliser une analyse des lacunes. Les résultats de cette analyse ont par la suite été validés par des entretiens avec la communauté et des groupes de discussion avec des représentants des bailleurs de fonds et des établissements.

Ces informations nous aident à la fois à définir les actions à entreprendre dans le cadre de la feuille de route et à donner la priorité à celles qui sont susceptibles d'avoir le plus de répercussions pour relever les défis concrets que la communauté de la recherche canadienne doit relever aujourd'hui.

3.3.1 Analyse des lacunes

Nous avons combiné les résultats des phases de consultation précédentes, qui avaient permis d'identifier les PID prioritaires, et exploré les priorités politiques du Canada en recourant à la fois à des recherches documentaires et à de nouvelles conclusions tirées d'entretiens, d'ateliers et de groupes de discussion, afin de dresser un tableau complet de l'activité en matière de PID et des progrès accomplis en vue de leur adoption. En commençant par cataloguer les acquis techniques, sociaux et politiques déjà connus, nous avons contextualisé les principales priorités de la communauté de recherche canadienne à partir des résultats de nos ateliers. Nous avons identifié les éléments de la stratégie sur les PID qui sont déjà en cours de réalisation ou de planification, et nous les avons mis en correspondance avec les priorités, les besoins et les exigences qui ne sont pas encore en place.

Nous avons examiné les composantes politiques, sociales et techniques en fonction de quatre thèmes :

1. Gouvernance
 - a. Technique
 - b. Sociale
 - c. Politique
2. Capacité technique
3. Leadership et clarté
 - a. Stratégie
 - b. Politique
 - c. Cadre
 - d. Priorités
4. Mobilisation et sensibilisation de la communauté

Le caractère dispersé de la stratégie sur les PID et de l'infrastructure nationale du Canada impose la mise en place d'une structure de gouvernance solide et soigneusement élaborée. Il faut veiller à ce que les diverses communautés soient bien représentées et que leurs besoins soient adéquatement pris en compte. Parallèlement, il convient de tenir compte de la nécessité d'assurer une normalisation adéquate pour garantir l'interopérabilité. Le renforcement de la gouvernance nécessitera la mise en place d'une structure plus formelle et devra disposer de ressources suffisantes.

Il faut élaborer une stratégie globale de gestion de la recherche qui s'appuie sur la vision, la mission et les principes. Un bon travail a déjà été accompli en ce sens et il faut continuer dans cette voie. Pour souligner la nécessité de consacrer davantage de ressources à la gouvernance, certaines décisions importantes doivent être prises de manière relativement urgente. Par exemple, le Canada souhaite-t-il se doter d'un système national de gestion de la recherche semblable à ceux mis en place au Portugal ou en République tchèque ? Quels sont les PID les plus importants pour les chercheurs, les établissements et les bailleurs de fonds au Canada ? Quels sont les besoins des différentes

communautés du Canada, par exemple les francophones ou les communautés autochtones, en matière d'infrastructures de recherche ?

Le travail actuel d'élaboration de politiques pour formuler l'énoncé du problème, l'argumentaire sur les PID et la note d'information doit être mis à profit pour sensibiliser les parties prenantes, les décideurs et les responsables de politiques importants au Canada. Cette démarche s'impose pour garantir une mobilisation des diverses communautés, sensibiliser les décideurs et les chefs de file, et débloquer des fonds. Il faut mettre en place une campagne de sensibilisation stratégique et un plan de mobilisation qui couvrent à la fois la sensibilisation de la base et l'adhésion des dirigeants.

À l'échelle des communautés, le travail des consortiums ORCID et DataCite a été très profitable. Il subsiste toutefois des lacunes sur le plan des connaissances, de la compréhension et des capacités au sein de la communauté canadienne de gestion de la recherche, notamment concernant l'importance et la valeur des PID ; il faut donc poursuivre les efforts visant à tirer parti de la réussite des deux consortiums. Par exemple, les systèmes d'information sur les recherches en cours (CRIS – current research information systems) sont des logiciels qui sous-tendent une grande partie du travail des gestionnaires de la recherche, en rassemblant les ressources au sein des universités et en s'intégrant aux PID et aux registres de métadonnées qui leur sont associés. La création de groupes d'utilisateurs de CRIS encouragera les communautés de pratique et aidera à définir les besoins en matière de gestion de la recherche au Canada. Cela renforcera à son tour la communauté des gestionnaires de recherche canadiens, leur permettant, par exemple, d'inciter les développeurs et les vendeurs de logiciels à mettre au point des fonctionnalités pertinentes et nécessaires.

Les bailleurs de fonds doivent montrer l'exemple en attribuant des identifiants de subvention et en faisant preuve d'un leadership politique plus visible dans des domaines tels que le changement technique et social. Les bailleurs de fonds devraient aussi intégrer ORCID et ROR dans leurs systèmes de financement. Ils exercent une grande influence sur tous les écosystèmes de recherche, en raison de leur capacité à assortir le financement d'exigences et à parrainer l'innovation. Ainsi, en adoptant des identifiants de subvention et en intégrant d'autres PID, et en faisant preuve d'un engagement et d'un leadership visibles, ils enverraient un message fort à la communauté.

Le développement des aspects techniques de la stratégie sur les PID au Canada sera très important pour le succès à long terme. Une fois qu'une stratégie sur les PID aura été élaborée et qu'une gouvernance solide aura été mise en place, il sera important de se pencher sur les capacités techniques afin que les établissements, les bailleurs de fonds et les autres parties prenantes soient en mesure de soutenir efficacement le développement et l'adoption des services de PID.

Un service d'assistance technique centralisé s'impose pour faciliter l'adoption des PID par les établissements et pour en réduire le coût et la charge pour chacune d'entre elles. Des recherches^{51,52} ont montré que des services de soutien technique centralisés pour aider les établissements dans les processus techniques, la formation et la gestion des changements liés à l'adoption des PID permettent de réduire le coût global de l'adoption à l'échelle systémique. Ces services sont également susceptibles d'accélérer le rythme d'adoption et de favoriser le développement de communautés de pratique, ce qui permettra d'améliorer l'interopérabilité et le partage d'informations.

Les modules d'extension des dépôts et des CRIS, les API et autres intégrations devraient bénéficier d'une aide financière centrale une fois que les besoins auront été identifiés. Pour que l'adoption des PID soit hautement efficace, des difficultés techniques doivent être surmontées. Comme de

⁵¹ J. Brown, P. Jones, A. Meadows et F. Murphy, « Revised cost-benefit analysis for the UK PID Support Network », Zenodo, nov. 2022. doi: 10.5281/ZENODO.7356219.

⁵² P. Jones, A. Meadows, F. Murphy, J. Brown, A. Conway et L. Griffith (2024). Efficiency and insight: a cost-benefit analysis for a central service to support persistent identifier implementation in Ireland. Digital Repository of Ireland. <https://doi.org/10.7486/DRI.nz80kt123>

nombreux établissements utilisent des systèmes logiciels semblables, il est possible de réaliser des économies d'échelle en fournissant un financement central pour concevoir les extensions logicielles nécessaires aux intégrations. Les extensions et les intégrations mises au point de manière centralisée améliorent également les niveaux d'interopérabilité et l'harmonisation des normes.

3.3.2 Établissements

Nous avons organisé deux groupes de discussion avec des représentants d'organismes de recherche canadiens, les 16 et 23 avril 2024. Le nombre de participants a été inférieur à ce que nous espérions (neuf personnes pour les deux réunions), mais nous avons obtenu des informations précieuses, notamment, comme l'a fait remarquer un participant, sur l'importance d'examiner le coût de renonciation – que se passera-t-il s'il n'y a pas de stratégie nationale ? Voici quelques-uns des autres grands thèmes abordés lors des deux réunions.

Aperçu

Les participants travaillent dans divers domaines, notamment la gestion des données de recherche, l'archivage numérique, la technologie des bibliothèques, les métadonnées/normes et l'analyse. Tous connaissaient assez bien, voire très bien, les PID. La plupart d'entre eux ont estimé que la technologie n'était pas un problème, mais que l'adhésion, tant au niveau des dirigeants que des utilisateurs, représentait un défi ; si certaines personnes sont très enthousiastes, d'autres sont totalement désintéressées. Cela se traduit par un manque de participation au sein de leurs établissements, ce qui rend le changement difficile. L'adhésion des groupes mal desservis est également un problème ; les participants ont noté que plus l'environnement est développé, moins il est probable qu'il offre des services aux groupes historiquement mal desservis (ou pas desservis du tout). Cela s'explique à la fois par le fait que les développements servent souvent des intérêts particuliers et que les réseaux existants et les utilisateurs de services sont généralement ceux qui « parlent le plus fort ».

Leadership

La plupart des participants ne connaissaient pas le CCCPID, mais tous ont perçu la nécessité de son rôle en matière de coordination centrale, de mise en commun de l'information, de regroupement, etc. Cela suggère qu'il faut de toute évidence mieux faire connaître le groupe et son rôle. Les participants ont également noté l'importance du leadership pour garantir que les investissements en matière de PID sont continus et durables, c'est-à-dire pérennes, ce qui est une fois de plus une occasion à saisir pour le CCCPID. Ils ont également reconnu la possibilité pour les organisations clés de jouer un rôle de leadership plus important. Par exemple, le soutien des trois agences constituerait une énorme incitation à l'adoption et à la mise en œuvre des PID par les chercheurs et les établissements. Comme l'a demandé un participant : « Si les trois agences ne le font pas, alors qui le fera ? »

Sensibilisation et soutien des communautés

Nous devons veiller à ce que la stratégie tienne compte, par exemple, des résultats non traditionnels, des locuteurs non anglophones, des besoins en connaissances autochtones, etc., et à ce qu'il y ait une proposition de valeur claire pour chaque groupe de parties prenantes. Un soutien et des conseils centralisés sont également considérés comme essentiels pour garantir la cohérence de l'utilisation et de la mise en œuvre des PID dans l'ensemble de la communauté de recherche au Canada. Cela passe par la formation et le perfectionnement des chercheurs, ainsi que par une documentation accessible – le tout en anglais et en français.

Préoccupations en matière de protection de la vie privée et de sécurité

Les participants aux deux réunions ont fait part de leurs préoccupations en matière de protection de la vie privée et de sécurité, notamment concernant RAiD et le partage d'informations à l'échelle des projets. « Tout système qui interagit avec d'autres systèmes devrait faire l'objet d'une vérification quant à sa sécurité technique. »

Qualité et fiabilité des données

Des inquiétudes ont également été exprimées quant à la fiabilité des données ; parmi les exemples cités figurent la fiabilité des informations contenues dans un dossier ORCID ainsi que la pourriture de liens.

Principaux PID

Les participants ne s'entendaient pas sur les PID qu'ils souhaiteraient privilégier, mais ORCID a été un favori incontestable et est déjà le plus utilisé. Les PID pour les PGD ont également été considérés comme une priorité élevée ou moyenne par de nombreux participants, mais les avis sur RAID étaient plus partagés, en partie parce que plusieurs participants ne le connaissaient pas. Les participants ont exprimé une préférence pour les PID émanant de la communauté, comme ROR. Ils ont mentionné certains identifiants qui ne figurent pas dans les priorités actuelles, notamment ARK (évoqué en relation avec l'engagement en faveur de la pérennité et l'accent mis par conséquent sur la qualité des [méta]données) ; NOID (Nice Opaque IDentifier), un producteur et un résolveur de noms pour les contenus numérisés ; et des identifiants pour les sujets/codes de sujet.

Systèmes et flux de travail

Les participants ont fait remarquer que beaucoup de leurs organisations n'ont pas de CRIS. Ils estiment aussi que les flux d'évaluation de la recherche devraient être considérés comme un élément important de la stratégie.

3.3.3 Bailleurs de fonds et agences

Il a été difficile par le passé de planifier des rencontres avec les bailleurs de fonds, et ces derniers s'inquiètent souvent d'un partage excessif des informations entre les agences. Nous avons donc recueilli les commentaires de ce groupe au cours de six entretiens (dont un avec deux personnes de la même organisation), menés entre la mi-avril et la fin avril. Collectivement, ils représentaient des bailleurs de fonds fédéraux et provinciaux, spécialisés dans une discipline donnée et généralistes au Canada.

Aperçu

La plupart des personnes interrogées ne connaissaient pas très bien la stratégie sur les PID et le projet de feuille de route du Canada, mais toutes comprenaient assez bien ce que sont les PID en général. Elles ont participé positivement aux discussions et se sont montrées disposées à envisager un avenir dans lequel les systèmes des bailleurs de fonds permettraient l'utilisation de PID.

Dans l'ensemble du secteur, les personnes interrogées ont clairement exprimé leur frustration face au fait que l'adoption des PID au sein de leur organisation a été plus faible et plus lente qu'elles ne l'auraient souhaité. Elles ont recensé un certain nombre de difficultés importantes qui devront être surmontées pour que des progrès significatifs puissent être accomplis à moyen terme.

Compréhension et utilisation actuelles des PID

L'utilisation des PID n'est pas optimale dans bon nombre de ces agences – même si des identifiants sont recueillis, ils ne sont pas encore utilisés pour relier et échanger des informations entre les systèmes. Les discussions internes sur les identifiants ORCID sont les plus avancées, mais plusieurs personnes interrogées ont indiqué que les identifiants ORCID pouvaient être ajoutés uniquement manuellement à leurs systèmes, sans autre forme d'interopérabilité permettant d'envoyer des informations vers les dossiers ORCID ou d'en recevoir. Des identifiants internes (et non des PID) sont utilisés pour les demandes de subvention, le suivi et l'établissement de rapports. Le FRQ fait exception : il a récemment mis en œuvre l'authentification unique d'ORCID, travaille actuellement à l'extraction d'informations à partir des dossiers ORCID et commencera ensuite à leur transmettre des informations.

Les personnes interrogées se sont intéressées à la plupart des PID mentionnés dans les entretiens et ont facilement compris la proposition de valeur qu'ils représentent. Elles ont exprimé un intérêt particulier pour les PID pour les plans de gestion des données et RAiD (même s'il faut préciser davantage la différence entre les identifiants de projet et les identifiants de subvention).

Les personnes interrogées spécialisées dans la technique et la gestion des connaissances étaient beaucoup plus familières avec les PID que les directeurs de programme et les cadres supérieurs. Le personnel technique est conscient de l'ampleur de la charge administrative et comprend comment l'introduction de flux de travail basés sur les PID dans les plateformes de gestion des subventions pourrait être un facteur de transformation. En revanche, les gestionnaires de programmes et les administrateurs sont très conscients des exigences gouvernementales en matière de traitement de la confidentialité et de l'hébergement des données, ainsi que des conflits d'intérêts ; ils sont sensibles aux pressions potentielles exercées par la communauté pour imposer l'utilisation obligatoire des PID aux bénéficiaires de subventions. Cette situation engendre une culture d'aversion au changement, avec des inquiétudes quant à l'utilisation des ressources de l'organisation pour investir dans de nouveaux systèmes.

Les identifiants internes pour les demandeurs et les bénéficiaires de subventions sont très courants. Leur utilisation précède l'identifiant ORCID, de sorte que certaines agences sont réticentes à passer à ORCID (ou aux PID en général). Les techniciens et les responsables des opérations ont du mal à convaincre leurs organisations d'intégrer les PID, car leurs collègues chargés des programmes, souvent plus haut placés, ne comprennent pas la proposition de valeur des PID et peuvent être réticents à se lancer en raison de leurs préoccupations en matière de conformité : l'introduction des PID au sein de ces organisations nécessitera une période de transition.

Le paysage des bailleurs de fonds

Toutes les personnes interrogées reconnaissent qu'il faudra résoudre les problèmes de ressources dans leurs organisations pour que des progrès puissent être réalisés. Il est difficile de prioriser les travaux de développement technique, et des incertitudes subsistent quant aux priorités futures des gouvernements ainsi qu'à la capacité de leurs organisations à obtenir le soutien à long terme dont elles auront besoin en matière de PID pour que le niveau d'investissement requis soit adéquat.

Plusieurs personnes interrogées ont déclaré que la Feuille de route pour la science ouverte⁵³ du gouvernement fédéral avait contribué à présenter des arguments internes en faveur des PID. Toutefois, il existe clairement une tension entre l'engagement national pour la science ouverte et les exigences prioritaires en matière de résidence des données, de protection de la vie privée et de la sécurité nationale, de licences et de conformité. Les bailleurs de fonds sont également très réticents à accepter des dépôts directs de données et à prendre des mesures visant à accroître l'ouverture qui pourraient accidentellement mettre en péril leur capacité à maintenir la confiance dans leurs autres responsabilités – d'autant plus que la gouvernance des PID échappe à leur contrôle direct.

Systèmes de gestion des subventions

Il règne un fouillis dans les systèmes de gestion des subventions de l'ensemble des bailleurs de fonds. Par exemple, les chercheurs qui travaillent pour les trois agences ont des comptes et des numéros d'identification différents pour chacune d'entre elles, qui ne sont pas reliés entre eux (les trois agences tentent toutefois de mettre en place un nouveau système). Des inquiétudes ont également été soulevées quant à la capacité à convaincre les développeurs de systèmes de subventions au service des bailleurs de fonds d'intégrer les PID.

⁵³ <https://science.gc.ca/site/science/fr/bureau-conseillere-scientifique-chef/science-ouverte/feuille-route-pour-science-ouverte>

Possibilités d'action et d'intervention

La participation des personnes interrogées a été très utile pour répertorier des possibilités de renforcer la sensibilisation, l'adoption, la pratique et la confiance des bailleurs de fonds à l'égard des PID. Elles sont regroupées par thème ci-dessous.

Conseils et directives documentés

Voici quelques suggestions de documents qui pourraient aider à plaider en faveur des PID, à susciter l'adhésion interne ou à apaiser les inquiétudes actuelles :

- Des énoncés clairs sur l'importance des PID du point de vue des bailleurs de fonds ;
- Des conseils sur la mise en œuvre, l'intégration, les meilleures pratiques pour les systèmes, les façons d'assurer la pérennité, etc. ;
 - Particulièrement sur une harmonisation mondiale avec d'autres bailleurs de fonds ;
 - Des conseils sur la manière de travailler avec les groupes d'utilisateurs fournisseurs ;
- Des indications explicites pour garantir une mobilisation équitable des communautés francophones ;
- Des conseils sur les priorités à établir, d'autant plus qu'il y a tant de choses à faire ;
- Des ressources pour réussir à convaincre les divers acteurs en interne ;
- Des exemples de réussite de gouvernements/d'agences, comme les documents australiens sur l'analyse coût-bénéfice ;
- Des éléments indiquant à quel point il serait plus facile d'obtenir des informations si les PID étaient mis en œuvre (p. ex., avec RAiD) ;
- L'importance d'identifier et de soutenir les défenseurs des PID.

Questions de gouvernement et de politique

Ces questions pourraient nécessiter plus de consultation ou une mobilisation en fonction du paysage législatif particulier des bailleurs de fonds et de leurs priorités :

- Les données doivent résider au Canada, ORCID ne peut donc pas être imposé. Les exigences en matière de sécurité (en réaction à une affaire récente de vol de données du gouvernement par des scientifiques chinois) ont été renforcées. Cette situation a compliqué l'environnement de prise de décision, mais elle pourrait offrir des possibilités de renforcer l'engagement des hauts responsables au sein des organismes de financement ;
- Déterminer comment permettre la prise en charge des CV narratifs et du CVC par le biais de flux de travail utilisant des PID ;
- Renforcer la confiance dans la possibilité d'une stratégie de sortie au besoin ;
- Mettre en évidence l'investissement plutôt que les coûts (dans ce cas, le coût serait : ce qui se passera s'ils ne prennent pas de mesures).

Collaborations

Quelques possibilités de mise en commun:

- Travailler avec des bibliothécaires et des organisations telles qu'Érudit et l'Acfas pour accroître la compréhension et l'adoption et, fondamentalement, garantir l'accessibilité par le biais d'une documentation non anglophone ;
- Collaborer avec les consortiums DataCite et ORCID pour accélérer la diffusion de messages via les réseaux existants (plusieurs personnes interrogées représentaient des bailleurs de fonds membres du groupe d'intérêt des bailleurs de fonds d'ORCID, qui s'est engagé à améliorer la coordination autour des PID⁵⁴).

⁵⁴ <https://info.orcid.org/wp-content/uploads/2024/02/ORCIDFigRecommendations2023.pdf>

3.3.4 Points importants des politiques

Les bailleurs de fonds interrogés ont établi des liens évidents entre l'orientation du gouvernement fédéral canadien et la vision interne de leurs agences quant à leur marge de manœuvre. Les bailleurs de fonds eux-mêmes ont également certaines priorités stratégiques de haut niveau, notamment des modifications au CV commun canadien (CVC). Un examen plus approfondi des points de contact qu'ils ont identifiés (ci-dessous) permettrait de comprendre s'il s'agit de catalyseurs ou plutôt d'obstacles potentiels au progrès, et de déterminer ce dont les différentes parties prenantes auront besoin pour aller de l'avant.

Les PID et les priorités actuelles des bailleurs de fonds

Tous les bailleurs de fonds que nous avons interrogés poursuivent des objectifs précis qui pourraient être intégrés dans l'argumentaire relatif aux PID, notamment la réduction de la charge de travail, le renforcement de l'accessibilité et le soutien au multilinguisme.

Le CVC est considéré par les chercheurs comme un mal nécessaire, notamment en raison de son interface utilisateur médiocre, du manque d'interopérabilité entre les agences et de la lourde charge administrative qu'il suppose. Toutefois, comme le CVC ne sera éventuellement plus maintenu par les trois agences, les mises à jour et l'interopérabilité pourraient enfin être possibles. Et d'autres bailleurs de fonds se prononceront sur la solution de remplacement, ce qui offre à nouveau la possibilité d'une option plus adaptée aux PID et moins contraignante.

Les PID et les priorités actuelles des gouvernements

Les personnes interrogées ont expliqué qu'une directive gouvernementale claire sur ce qu'il convient de faire au sujet des PID serait utile pour débloquer et orienter les ressources, ainsi que pour renforcer la confiance dans les systèmes de PID. Une réponse officielle ou un manuel expliquant les points de recoupement entre les PID et la nécessité de clarifier la question des licences, de gérer les différentes exigences en matière de confidentialité et de garantir la conformité en matière de conflits d'intérêts et de sécurité des données, auraient un impact considérable. Il serait encore plus utile de mettre en relief la façon dont les PID peuvent relier la recherche et les chercheurs canadiens à la communauté internationale.

Le gouvernement canadien amorce le déploiement de normes en matière de référence des données (à partir de décembre 2023, en commençant par une norme pour les provinces et territoires canadiens). Ces normes incluront probablement des organisations, ce qui permettra de mettre en correspondance les identifiants organisationnels du gouvernement avec les identifiants ROR. Un examen plus approfondi de ces normes pourrait permettre de garantir que, dans la mesure du possible, des PID soient utilisés pour minimiser le travail et la complexité, par exemple en réduisant le plus possible le nombre de passerelles entre les systèmes.

Plusieurs personnes interrogées ont indiqué que la Feuille de route pour la science ouverte⁵⁵ du gouvernement canadien pourrait s'avérer utile. Pour atteindre l'objectif global de la science ouverte, et en tenant compte des réactions émanant des consultations intraministérielles (recommandation n° 2), les agences et les ministères ont été chargés de fournir des plans d'action⁵⁶ indiquant comment ils prévoient de mettre en œuvre les recommandations de la Feuille de route pour la science ouverte avant juin 2021. Tous les bailleurs de fonds n'ont pas publié de tels plans, mais le CNRC⁵⁷ l'a fait.

⁵⁵ <https://science.gc.ca/site/science/fr/bureau-conseillere-scientifique-chef/science-ouverte/feuille-route-pour-science-ouverte>

⁵⁶ <https://science.gc.ca/site/science/fr/science-ouverte/plans-daction-miniserials-pour-science-ouverte>

⁵⁷ <https://nrc.canada.ca/fr/organisation/planification-rapports/plan-daction-science-ouverte-reponse-feuille-route-science-ouverte-gouvernement-canada>

4. Recommandations

Les recommandations présentées ci-dessous sont classées en fonction de la structure de l'analyse des lacunes. Ces catégories sont utiles ici dans la mesure où elles indiquent les types d'expertise et d'activités de projet qui sont nécessaires pour mettre en œuvre chaque recommandation. Elles pourraient servir à structurer la conception du programme en vue de la mise en œuvre de la Feuille de route.

4.1 Leadership et clarté

Il convient de mettre en place un leadership fort et visible. La nature distribuée de l'infrastructure des PID présente des risques pour la dynamique, que seules une mobilisation et une coordination proactives permettront de surmonter. Si le CCCPID a fait et continue de faire un excellent travail pour maintenir en vie le programme de PID, son rôle actuel est explicitement consultatif, sans mandat formel de leadership.

Pour que les PID soient adoptés, il faut que plusieurs groupes distincts agissent en harmonie. Et comme les avantages des PID sont générés en grande partie par des effets de réseau, chaque groupe a besoin de l'engagement des autres pour pouvoir tirer la valeur souhaitée de ses investissements dans l'intégration et l'adoption des PID. C'est là une problématique classique des actions collectives, qui nécessitera une forte impulsion politique centrale, idéalement au niveau de la conseillère scientifique en chef ou du gouvernement fédéral, afin de hisser ce programme au sommet de la liste des priorités organisationnelles concurrentes. Sans ce niveau d'« appropriation » de la stratégie, il ne sera pas possible de débloquer les types d'investissements nécessaires à un accès cohérent et équitable aux PID. La coordination intercommunautaire sera alors beaucoup plus difficile à réaliser et, ultimement, il n'y aura pas grand-chose pour faire passer l'intégration des PID en tête de la liste des choses à faire, étant donné les nombreuses priorités concurrentes dans le secteur.

4.2 Gouvernance

Il est urgent d'établir une feuille de route pour mettre en place une gouvernance technique, sociale et politique solide. Il faut prendre un certain nombre de décisions essentielles quant à la manière dont le Canada souhaite traiter les informations relatives à la gestion de la recherche, et il est impératif de consulter les communautés traditionnellement mal desservies au Canada, telles que les peuples autochtones et les communautés francophones.

En se fondant sur des décisions structurelles, techniques et de gouvernance, il faut établir un ordre de priorité pour déterminer les PID qui devraient être les plus fortement soutenus, tout d'abord sur la base des consultations de la communauté et de l'avis d'un comité consultatif technique. Cette priorisation doit être considérée comme un document évolutif et être constamment révisée en fonction de l'évolution des besoins stratégiques du Canada en matière de gestion de la recherche.

4.3 Capacités techniques

Un certain nombre de projets techniques pourraient contribuer à favoriser l'adoption et l'utilité des PID. Bon nombre de ces programmes potentiels, par exemple les modules d'extension des dépôts pour des PID donnés, sont tributaires des décisions de gouvernance technique et sociale décrites ci-dessus. Dans l'intervalle, il conviendrait de mettre en place un service d'assistance technique central pour l'adoption des PID.

Les discussions avec la communauté ont mis en évidence plusieurs intégrations de PID à privilégier. Nous les avons réparties en trois catégories générales, car elles permettent d'illustrer les parties prenantes et les systèmes concernés par chacune d'entre elles.

Apports à la recherche

Les bailleurs de fonds canadiens devraient montrer l'exemple et commencer dès que possible à délivrer des PID pour les subventions (par opposition aux numéros de subvention internes existants). Cette fonctionnalité devrait être intégrée dans les systèmes de gestion des subventions et des bourses. Il est vivement recommandé de l'intégrer ensuite dans les systèmes d'établissement de rapports et les bases de données internes de la recherche, afin de garantir que la valeur de ces PID se traduise par un retour sur investissement pour les bailleurs de fonds et de maximiser l'allègement de la charge administrative pour les chercheurs.

Activités de recherche

Le Canada devrait dès que possible adopter des identifiants d'activités de recherche (RAiD) pour les projets, en sachant que le système est à un stade avancé de développement et que, au moment de la rédaction du présent document, son déploiement mondial est prévu pour le second semestre de 2024. Entretemps, un sous-groupe du CCCPID devrait étudier les options et les exigences d'une agence d'enregistrement RAiD pour le Canada ou évaluer un partenariat avec l'une des autres agences d'enregistrement potentielles d'Amérique du Nord. Ces PID présentent un avantage supplémentaire : la transparence qu'ils pourraient apporter aux activités et aux collaborations mondiales.

Accroître le financement et les capacités du consortium ORCID-CA, éventuellement en recrutant du personnel supplémentaire par l'intermédiaire du service de soutien en matière de PID mentionné ci-dessus, avec pour mission d'améliorer les intégrations et la cohérence d'ORCID, et de se concentrer sur l'alimentation des dossiers ORCID avec des données fiables. Cela implique en premier lieu des intégrations dans les systèmes des établissements et des bailleurs de fonds, avec la capacité d'écrire dans les dossiers ORCID et d'ingérer des données à partir de ceux-ci.

Il convient de mettre en place des identifiants communs, mondiaux et ouverts pour les organismes. Ils devraient être fortement encouragés dans les processus de gestion de la recherche, et éventuellement rendus obligatoires pour l'établissement de rapports. Les identifiants ROR sont actuellement les plus utilisés et devraient être évalués en vue de leur utilisation par les organismes de recherche, dont les bailleurs de fonds, et en tenant compte de leur utilisation potentielle pour l'identification des installations de recherche.

Il nous faut de toute urgence des PID pour les instruments, comme les DOI, pour suivre et gérer les investissements et garantir l'intégrité de la recherche. Cela nécessitera des intégrations dans les systèmes institutionnels, à la fois pour la réalisation de la recherche, dont les cahiers de laboratoire électroniques, et pour les systèmes de gestion de la recherche.

Résultats de la recherche

Tout au long de l'analyse contextuelle, les PID pour les ensembles de données, principalement les ARK pour les travaux en cours et les DOI pour les documents statiques ou publiés, sont ressortis des discussions comme une priorité pour les établissements et les praticiens. Comme pour les instruments, cela suppose des intégrations dans les systèmes institutionnels, mais avec une valeur supplémentaire à tirer des intégrations par les bailleurs de fonds dans les systèmes de rapport et d'évaluation. Cela permettra également de suivre les progrès accomplis dans la réalisation des objectifs politiques en matière de disponibilité, de souveraineté et de réutilisation des données.

Il faudrait vivement inciter tous les éditeurs canadiens de publications savantes à utiliser des DOI pour les produits publiés. Sachant que les publications plus petites, émergentes ou non traditionnelles pourraient ne pas disposer des ressources ou de l'équipement nécessaires pour accéder à ces DOI ou en tirer parti, le CCCPID devrait étudier les besoins des petits éditeurs et la possibilité d'un parrainage ou d'un mécanisme comparable en vue d'améliorer l'accès et la couverture.

Le fil conducteur de toutes ces actions est la nécessité pour les établissements et les bailleurs de fonds de prioriser dans la mesure du possible la réutilisation des métadonnées des registres de PID, afin de réduire au minimum la duplication des efforts et les erreurs résultant de la saisie manuelle des données.

4.4 Mobilisation et sensibilisation de la communauté

Les PID et les infrastructures de recherche doivent occuper une place plus importante sur la liste des priorités politiques nationales. Plusieurs produits issus de ce travail et de travaux antérieurs, notamment l'énoncé du problème, l'argumentaire sur les PID et la note d'information, peuvent constituer des éléments centraux de la communication. Une stratégie de sensibilisation s'impose pour tirer le meilleur parti de ces ressources. Pour garantir la mobilisation des francophones, une stratégie sur mesure pourrait inclure des collaborations avec d'autres parties prenantes, telles qu'Érudit et l'Acfas. Il pourrait être utile de traduire en français d'autres productions de MoreBrains relatives aux PID, qui ne sont pas directement liées à ce projet et sont disponibles en libre accès.

Lors de la préparation du contenu de la sensibilisation, il convient de veiller à ce que les documents expliquent où se situe la valeur de l'intégration des PID. On pourrait, par exemple, insister davantage sur le fait que la valeur ajoutée des identifiants ORCID réside dans l'augmentation du nombre d'intégrations et le temps gagné grâce à la réduction de la charge administrative, plutôt que dans le nombre des identifiants ORCID enregistrés.

Des travaux devront être menés afin d'aborder directement la situation des bailleurs de fonds qui ignorent encore quels seront les effets de l'introduction des PID sur leurs responsabilités générales en matière de sécurité des données, de respect de la vie privée, de résidence des données, etc. Au besoin, il faudrait miser sur le soutien de l'initiative de la Feuille de route pour la science ouverte. Les bailleurs de fonds sont actuellement pris en étau entre ces deux impératifs, de sorte qu'un aperçu de la manière dont les PID pourraient aider à les gérer pourrait se révéler transformateur.

Un soutien accru de la part de la communauté s'impose. Les consortiums ORCID et DataCite ont fait un excellent travail de coordination entre les différents acteurs institutionnels. En renforçant leur capacité par un soutien supplémentaire et en assurant la visibilité des activités de la communauté par une coordination entre les parties prenantes – via par exemple les groupes d'utilisateurs de CRIS et de dépôts, les associations professionnelles nationales –, on contribuera à maintenir l'élan et à diffuser les bonnes pratiques.

Pour motiver et maintenir le mouvement, il faudra disposer de preuves du retour potentiel sur investissement et en contrôler la réalisation concrète et effective. On pourrait notamment adapter au public canadien des analyses existantes, comme les analyses coûts-avantages réalisées par MoreBrains au Royaume-Uni⁵⁸, en Australie⁵⁹ et en Irlande⁶⁰, et offrir ainsi des approches méthodologiques utiles pour évaluer l'ampleur et le taux d'adoption.

⁵⁸ J. Brown, P. Jones, A. Meadows et F. Murphy (2022). Revised cost-benefit analysis for the UK PID Support Network (Version 1). Zenodo. <https://doi.org/10.5281/zenodo.7356219>

⁵⁹ J. Brown, P. Jones, A. Meadows et F. Murphy (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>

⁶⁰ P. Jones, A. Meadows, F. Murphy, J. Brown, A. Conway et L. Griffith (2024). Efficiency and insight: a cost-benefit analysis for a central service to support persistent identifier implementation in Ireland. Digital Repository of Ireland. <https://doi.org/10.7486/DRI.nz80kt123>