

MORE+BRAINS

Towards a national PID strategy for Canada

CPIDAC PID consultancy final report



Le français suit

30th May 2022

MORE+BRAINS

Executive Summary

Canada benefits from a lot of energy and engagement within a core persistent identifier (PID) community. The Canadian research community contains considerable expertise in the practical and policy issues relevant to PIDs, and is well connected with other PID consortia, experts, initiatives, and providers globally. It is steered by the Canadian Persistent Identifier Advisory Committee (CPIDAC), which brings together key stakeholders and experts in articulating a vision of a research system in which PIDs deliver:

- administrative efficiencies and reduced bureaucratic burden;
- enhanced information systems that provide greater insight, accountability, and impact for Canadian research;
- better integration between institutional, local, provincial, national, and international systems in the service of improving the targeting and strategic management of research support;
- opportunities for Canada to lead the world in the development and implementation of crucial research technologies.

This project began with a landscape review of PID advocacy and adoption in Canada, complemented by an analysis of national PID policies and activities in other countries (such as Australia, Finland, the Netherlands, and the UK). Informed by this desk research, the project team conducted six interviews with Canadian experts from a range of professional arenas to gather their insights into the factors accelerating and hindering PID adoption and use in Canada.

Finally, the team, working closely with staff from the Canadian Research Knowledge Network (CRKN), and the Digital Research Alliance of Canada (the Alliance), facilitated three virtual consultation workshops, which brought together a diverse group of 40 stakeholders representing a broad swath of the Canadian research community. The findings and outcomes from these workshops gave both depth and breadth to an assessment of the current state of PID adoption in Canada.

It is clear that there are real challenges to be overcome on the path toward a PID-optimized ecosystem. These include factors such as differing levels of trust between some stakeholder groups; fragmented infrastructure across funding organizations, which can make it hard for researchers to understand what is expected of them; and a lack of urgency in the broader community, who see PIDs (to the extent that they are considered at all) as opaque, overly technocratic, or subject to projection of concerns and mistrust felt towards policies or authorities.

To overcome these challenges, the benefits of PID adoption and integration must be both tangible and effectively communicated. [Section 3.1](#) provides details of community priorities for PID integrations. Leadership is both needed and desired to build consensus and maintain program focus: a broad coalition to coordinate the design and implementation of the PID strategy, evaluating options and thinking about the fit of particular interventions to wider goals; and a figurehead organization for this work with sufficient political and cultural capital to help to drive this agenda forward, especially at the Federal level.

PID adoption needs to be accelerated within a cohesive policy and support framework that can balance competing incentives, and help to ensure that practical assistance and concrete advantages of adoption are distributed equitably.

This report recommends that the CPIDAC should:

1. Build a grand coalition, broadening its reach and remit to bring in a wider range of stakeholders (e.g. maximize participation from professional groups and a variety of research institutions, or formalize researcher representation).
2. Articulate the value proposition for PIDs for those groups who need to act or engage (without input from the 'grand coalition,' a mismatch in aspirations could slow or prevent progress).
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.
5. Provide financial, technical, and social resources to support the development and governance of a common framework for PID access and data re-use (the technical and social components of this are inseparable, think of resilience, trust etc.).
6. Identify short-, medium-, and long-term actions to be taken en route to delivering the vision, and get community, sector, or agency-specific commitments to undertake those actions.
7. Take ownership of the PID agenda in Canada and consult, communicate, then consult again to ensure buy-in for the strategy.

In more concrete terms, the immediate next step is to set out a clear roadmap for the creation of a longer-term national PID strategy, including:

- establish a vision, co-created with a broad range of stakeholders, for what the Canadian research information landscape will look like in 5-10 years time. This will form the core of the roadmap;
- within this vision, prioritize a core set of PIDs and the use cases they would serve;
- work backwards from the vision to identify the preconditions for success (perhaps through a 'pre-mortem' exercise and/or other forms of scenario planning to identify risks or confounding factors);
- focus resources on moving the needle on adoption and coverage for the core PIDs (and associated metadata);
- extend the evidence base and governance framework to ensure political and financial support for delivery of the strategy.

Table of Contents

1 Background	4
2 Approach	4
3 Discussion	6
3.1 Overall findings	6
3.2 Considerations	8
4 Recommendations	10

1 Background

This report provides a detailed view of a project to map out the steps necessary for the creation of a Persistent Identifier (PID) strategy for the Canadian research community. It builds on years of successful PID advocacy within Canada, the establishment of two major PID consortia (providing access to, and technical support for, PIDs from DataCite and ORCID, covering outputs and researchers respectively), and the creation of the Canadian Persistent Identifier Advisory Committee (CPIDAC). It highlights the opportunities and challenges that should form the focus of future activities, and offers practical recommendations on ways these could be prioritized and addressed by an extended 'grand coalition' of Canadian research stakeholders.

In response to the CPIDAC request for proposals for a consultant to develop a national PID strategy for Canada, MoreBrains Cooperative put forward a successful proposal that envisaged a multi-step process for creating such a strategy: a roadmap to a national PID strategy for Canada.

The resulting project is based on the principle that advances in Canada's PID infrastructure will depend on a consensus being reached on suitable standards, workflows, best practices, and governance approaches across a range of Canadian research stakeholders. Each national research environment is unique as it is shaped by history and cultural norms, as well as current incentives and research strategy priorities. Establishing genuine buy-in and commitment to the principles underpinning the PID strategy will take time and consultation, and the strategy itself must not only be carefully researched, but also co-created and socialized with Canadian funders, institutions, and publishers.

Since November 2021, the MoreBrains team, supported by CPIDAC members, have worked with a diverse set of stakeholders to establish the requirements and scope for a robust and pragmatic Canadian PID strategy. This report sets out the findings of our community consultation, which had the following goals:

- identify and describe strategic research objectives;
- highlight potential priority PIDs required to meet those objectives, with recommendations for any additional research needed to assess adoption and awareness levels for these PIDs;
- recommend workflows and best practices to ensure adequate connectivity of PIDs through metadata;
- engage with the wider community to socialize the need for a PID strategy;
- identify unmet technical and governance needs;
- identify potential existing organizations or stakeholder groups to take on those responsibilities and/or any new structures or organizations needed;
- indicate governance, participation, and outreach strategies for any new or existing structures or organizations.

2 Approach

The MoreBrains team began with desk research, exploring a range of documents and resources prepared by the CPIDAC, existing Canadian PID consortia and various initiatives associated with them, as well as documentation from PID strategies and services internationally. This was followed

by a series of semi-structured interviews, the findings of which were used to inform the design of three stakeholder workshops.

This approach enabled us to explore the landscape more broadly, and then home in on specific factors and considerations to be investigated.

Interviewees were selected by a group from the Canadian ORCID and DataCite consortium leads at the Canadian Research Knowledge Network (CRKN) and the RDM team of the Digital Research Alliance of Canada (the Alliance), and were conducted by Josh Brown, MoreBrains co-founder and lead for research and strategy. The six interviewees represented a range of perspectives, from broad categories made up of:

- Committees/Institutions/Members
- Policy
- Funders
- Researchers

Some interviewees have been, or are, involved in the CPIDAC or one of the existing Canadian PID consortia governance groups, while others are less directly engaged with PIDs.

Findings from the interviews informed a series of three 'by invitation' stakeholder workshops, held virtually between February 28 and March 15, 2022, in which questions raised in the interviews were explored with larger groups of stakeholders drawn from across the Canadian research community. Workshop participants were identified and selected by the project sponsors, with feedback and recommendations from the MoreBrains team.

There were 40 participants from across the Canadian research community overall, drawn from a range of stakeholder groups, including provincial/national funders, government, universities (including libraries, research offices, IT), healthcare, publishing, and researchers.

The original six interviewees were invited to participate in the workshop series, and were spread across the three sessions. The third workshop included a mix of some delegates from workshops one and two plus newcomers, to ensure a good range of perspectives on the issues raised.

The project sponsors attended all workshops – primarily as observers, although some individuals with particular expertise (strategy, implementation) participated formally in the relevant session.

The workshops were conducted over Zoom, and supported with materials, presentations etc. in both English and French. CRKN staff were on hand to provide simultaneous translation when needed.

Plenary discussions were stimulated and structured using online polling on Mentimeter.

What pain points do you see around information sharing, re-use, or exchange in the sector at the moment?

Mentimeter



Figure one: Example of a Mentimeter poll from the workshops (NB: only English language shown in this screenshot)

Breakout sessions took deeper dives into specific issues in a virtual whiteboard environment, using Miro.

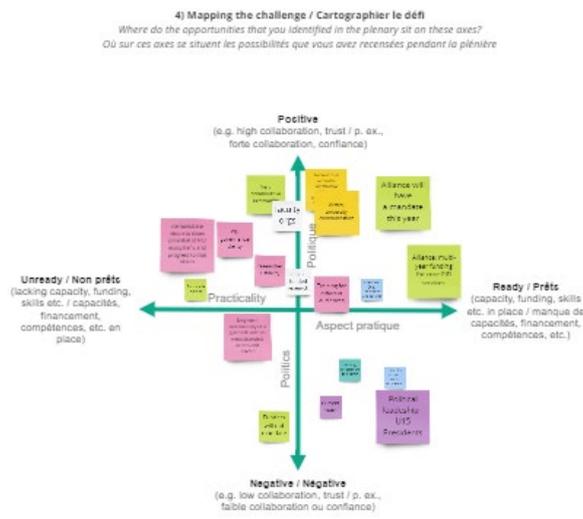
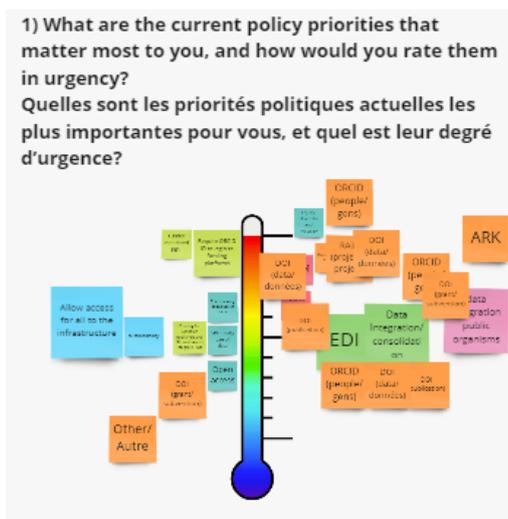


Figure two: Examples of breakout materials from virtual whiteboard sessions

These discussions, complemented by desk research and the MoreBrains team’s experience of conducting similar projects in other countries, were used to shape our recommendations to the CPIDAC on subsequent steps towards the development and launch of the Canadian PID implementation strategy.

3 Discussion

3.1 Overall findings

This section presents the main lessons and findings of the research phase of the project.

Some of these issues are mitigated, to an extent, by the prime assets identified as already in place to support the evolving Canadian PID strategy:

- there is a lot of energy and engagement within a core PID community;
- the Canadian research community benefits from existing expertise in the practical and policy issues relevant to PIDs, and is well networked with other PID consortia, experts, initiatives, and providers globally;
- Canada has an opportunity to take a leading role in balancing open data (etc.) and data sovereignty (FAIR and CARE). While breaking new ground comes with some risks, the risk/reward balance will be favourable if the PID strategy is inclusive by design and bolstered by practical and financial support to enable full participation from all communities.

Our analysis of the three workshops in particular surfaced a number of factors pertinent to the early stages of developing a national PID strategy for Canada:

- there are differing levels of trust between some stakeholder groups. This may limit buy-in at first, and will require consultation and ongoing communication to overcome;
- infrastructure is heavily fragmented across funding organizations. This makes it hard for researchers to understand what is expected of them, increases costs for service providers, and makes it more challenging to make a business case for the adoption or use of any particular PID;
- there is a lack of urgency in the broader community, who see PIDs (to the extent that they are considered at all) as opaque, overly technocratic, or subject to projection of concerns and mistrust felt towards policies or authorities;
- there is inconsistent understanding and acceptance of ORCID and other PIDs across and within communities (see preceding bullets), which ultimately can only be addressed with consistent and carefully designed messaging and education.

The workshops and interviews both emphasized a common set of critical insights:

Benefits must be both tangible and effectively communicated.

Funders and research institutions must ensure that researchers see benefits, not new burdens, as a result of this work. Since there may be some new burdens during transition (e.g. the cognitive load of learning a new way of undertaking a task, or in fully populating an ORCID record), or refocused burdens afterwards (where there may still be work to be done but with the emphasis shifted, for example, from providing basic factual information to providing context to automatically gathered factual, information), it may be more realistic to say that the benefits must be seen to outweigh the burdens.

The priorities highlighted in the workshops can help identify pathways to demonstrating these benefits, including:

- integrations into institutional systems;
- funder systems with better and (most importantly) consistent PID integrations;
 - Two-way flow to ORCID, etc.;
 - Registering grant funding IDs;

- data-sharing solutions specific to Canada (i.e. Canada-based data storage) to simplify links to PID registries and increase data sharing between research institutions and funders.

Leadership is both needed and desired.

In practice, at this stage, 'leadership' means driving a process of consensus-building, and maintaining program focus. There is clearly a need and a desire for a broad coalition to coordinate the design and implementation of the PID strategy, evaluating options and thinking about the fit of particular interventions to wider goals, and a figurehead organization for this work with sufficient political and cultural capital to help to drive this agenda forward, especially at the Federal level.

The workshop discussions of funder-issued mandates for PIDs provide an interesting example. Historically, mandates have boosted coverage, as in the Wellcome Trust requirement for grant applicants to provide an ORCID ID. However, they have also had unintended consequences, for example, the number of empty ORCID records created largely to comply with a mandate and then abandoned. Questions to consider in terms of mandates include: Is this a one-size-fits-all approach? Which would work better, endorsement or mandate? Since mandates can take many forms (e.g. requiring grantees to cite a Grant DOI in funding acknowledgements, or requiring repositories to register DOIs for all outputs of a particular type), it makes sense to tie them to the specific adoption of functional goals of the wider strategy. Given the risk of 'blowback' from mandates, a great deal of community consensus-building is needed, with a public and open weighing of benefits and implications.

The first task emerging from the workshops, then, is to establish a leadership for this process with the widest consultation possible. Once this is in place, the priority should be reaching agreement on a strategy before deciding on specific tactics, such as mandates.

3.2 Considerations

Across the research process, we have identified three significant themes which have a major bearing on the next steps in developing a national PID strategy for the Canadian research community.

Adoption

For PIDs and their associated metadata to be reliably useful, they need to be present and available to use for the majority of the entities they identify. The coverage of any given PID system is the key to its utility. PIDs need to be registered and shared at the earliest possible stage in the lifecycle of an entity, which means that the tools used to record or create those entities must have highly functional PID integrations. At the same time, the metadata associated with them needs to be accessible in structured, reusable forms for integration in administrative, discovery, reporting, and analytics systems.

At present, the benefits of PIDs are largely abstract and will remain so until PIDs and their associated metadata are used and reused predictably and consistently in systems, platforms, and tools across the Canadian research network.

Our research for this project showed that, at present, there are simply not enough technical integrations of PIDs. As a result, researchers have to manually add the same information into multiple funding and/or research institutional systems, multiple times.

We heard from research participants that a major barrier to adoption is the cost and complexity of PID integration. Many smaller research institutions cannot afford the technical infrastructure needed, and unique 'home-grown' implementations imply a duplication of effort and, therefore, a higher cost than the sector has capacity for. At the same time, participants reported that many commercial systems which offer PID integrations (e.g. CRISs) are too expensive for a number of research institutions.

One suggestion that came up in several workshop discussions was to create a Canadian information-sharing solution, hosted locally, to support access to PID APIs and metadata, combined with other information. This could take the form of a metadata store that can connect to PID APIs, funder systems, and institutions alike. This idea was originally suggested in workshop two (continuing a similar discussion which has been under way for some time¹), and then endorsed in workshop three. It could be a valuable way to lower costs, drive adoption, and ensure equitable access to PID APIs and services.

Cohesion

The absence of a coherent and consistent approach to PIDs within and across stakeholder groups is a major gap that needs to be filled, which is why the CPIDAC commissioned this work. While, as noted above, there is a core group of highly motivated individuals engaged with this initiative, the reach of this group needs to grow for the next stages in the process of 'PID optimization' to be delivered efficiently and effectively. The adoption of PIDs is currently hampered by the lack of a unified national vision and goals, as seen in policy variations across funders, inconsistent incentives, and a lack of clear policy priorities. These sector misalignments have become roadblocks to an efficient and PID-optimized research landscape.

Reconfiguring and aligning priorities and activities is therefore necessary for the PID strategy to succeed. Delivering this will require tackling historical issues, and fostering constructive engagement with the potential for PIDs to address sources of frustration for the research community. One example that came up many times in interviews and workshops was the Canadian Common CV. A number of participants told us that the CCV creates administrative burden, and it has clearly caused some deep scars for many of those who use it.

These issues may prevent some stakeholder groups from engaging with the PID strategy process, or from acting in a timely manner. In discussions, some funders expressed concerns about moving forward with solutions due to reputational risk. A cohesive, open, and inclusive prioritization process will help make the business case for such work, ensuring that priorities are shared and supported by a plurality of stakeholders and helping mitigate their concerns.

¹ Barsky, E. (2020, December 14). Research Information Management Integration for Canada [R]. doi:<http://dx.doi.org/10.14288/1.0395275> "We propose that NDRIO engage in a leadership role in implementing a national-scale, lightweight, interoperable RIM integration for Canada as a vital component of the overall RDM ecosystem."

Equity

A pertinent lesson from other national PID initiatives is that no group, organization, or community will receive the full benefits of PIDs until all do². The network effects and economies of scale that drive the benefits rely on participation that is as complete as possible.

There are many dimensions to this: small versus large institutions; local, provincial, and national scales; English- and French-speaking communities; Indigenous scholarship; disciplinary differences; and more. Other countries that are exploring PID strategies have similar challenges, so there is much common ground and experience upon which to build. Canada has a unique blend of cultural differences and multiple languages; while some aspects are analogous to other countries, such as Australia, its history and context are unquestionably distinctive.

Discussions in interviews and workshops surfaced some of the ways that language creates challenges at multiple levels. Language can denote cultural groups, it can raise barriers to access via monolingual interfaces, and it can also complicate the reuse of metadata or information.

Other cultural and political challenges that must be addressed directly include the risks of paternalism, appropriation, or under-representation of Indigenous peoples, and concerns around privacy and quantification (offered as a hypothesis to explain the slow adoption of ORCID IDs, particularly in Francophone communities).

Building bridges between groups, cultures, and communities will be essential if the PID strategy is to succeed, and all who participate in or contribute to Canadian research are to benefit from it. While the challenges are specific to Canada, there is an opportunity to lead by example on a global level – but this will require a transparently and inclusively developed model of data sovereignty.

4 Recommendations

In conclusion, we recommend that the CPIDAC should:

1. Build a grand coalition, broadening its reach and remit to bring in a wider range of stakeholders (e.g. maximize participation from professional groups and a variety of research institutions, formalize researcher representation).
2. Articulate the value proposition for PIDs for all groups who need to act or engage (without input from the 'grand coalition' a mismatch in aspirations could slow or prevent progress).
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.
5. Provide financial, technical, and social resources to support the development and governance of a common framework for PID access and data re-use (the technical and social components of this are inseparable, think of resilience, trust etc.).
6. Identify short-, medium-, and long-term actions to be taken en route to delivering the vision, and get community, sector, or agency-specific commitments to undertake those actions.

² See, for example: Brown, Josh, Jones, Phill, Meadows, Alice, Murphy, Fiona, & Clayton, Paul. (2021). UK PID Consortium: Cost-Benefit Analysis (1.0). Zenodo. <https://doi.org/10.5281/zenodo.4772627>

7. Take ownership of the PID agenda in Canada and consult, communicate, then consult again to ensure buy-in for the strategy.

In more concrete terms, the immediate next step is to set out a clear roadmap for the creation of a longer-term national PID strategy, including:

- establishing a vision, co-created with a broad range of stakeholders, for what the Canadian research information landscape will look like in 5-10 years time – this will form the core of the roadmap;
- within this vision, prioritize a core set of PIDs and the use cases they would serve;
- work backwards from the vision to identify the preconditions for success (perhaps through a ‘pre-mortem’ exercise and/or other forms of scenario planning to identify risks or confounding factors);
- focus resources on moving the needle on adoption and coverage for the core PIDs (and associated metadata);
- extend the evidence base and governance framework to ensure political and financial support for delivery of the strategy.

Vers une stratégie nationale sur les PID pour le Canada

Rapport final de la consultation sur les PID
réalisée pour le CCCPID



30 mai 2022

Sommaire

Le Canada peut compter sur un grand dynamisme et une importante mobilisation au sein d'une fidèle communauté d'utilisateurs des identifiants pérennes (PID). La communauté de recherche canadienne possède une grande expertise dans les questions pratiques et politiques relatives aux PID, et elle entretient des liens étroits avec d'autres consortiums, expertes et experts et fournisseurs de PID à l'échelle mondiale. Elle est représentée par le Comité consultatif canadien sur les identifiants pérennes (CCCPI), qui réunit des intervenantes et intervenants et des experts de premier plan dans le but d'élaborer une stratégie pour établir un système de recherche dans lequel les PID offrent :

- des gains d'efficacité et une diminution du fardeau administratif;
- des systèmes d'information améliorés qui offriront une meilleure compréhension du milieu de la recherche canadienne, de même qu'une reddition de comptes et mesure des retombées bonifiées;
- une meilleure intégration entre les différents systèmes (institutionnels, provinciaux, nationaux et internationaux) permettant de mieux cibler le soutien à la recherche et sa gestion stratégique;
- des occasions pour le Canada d'être un chef de file mondial dans la conception et la mise en œuvre de technologies de recherche de premier plan.

Pour ce projet, un portrait de la promotion et de l'adoption des PID au Canada a d'abord été établi pour ensuite être complété par une analyse des politiques et des activités nationales en matière de PID dans d'autres pays (comme l'Australie, la Finlande, les Pays-Bas et le Royaume-Uni). Grâce à cette recherche documentaire, l'équipe de projet a mené six entretiens avec des expertes et experts canadiens provenant de divers milieux professionnels afin de recueillir leurs points de vue sur les facteurs qui accélèrent et entravent l'adoption et l'utilisation des PID au Canada.

Enfin, en étroite collaboration avec des employés du Réseau canadien de documentation pour la recherche (RCDR) et de l'Alliance de recherche numérique du Canada (l'Alliance), l'équipe a animé trois ateliers de consultation virtuels, qui ont réuni un groupe diversifié de 40 acteurs représentant un large éventail du milieu canadien de la recherche. Les conclusions et les résultats de ces ateliers ont permis à la fois d'élargir et d'approfondir l'évaluation de l'état actuel de l'adoption des PID au Canada.

Il existe manifestement de véritables défis à relever sur la voie vers un écosystème optimisé par l'utilisation des PID. Il s'agit notamment de facteurs comme les différents niveaux de confiance entre certains groupes de parties prenantes; une infrastructure fragmentée entre les organismes de financement, ce qui peut faire en sorte que les chercheuses et chercheurs ont du mal à comprendre ce que l'on attend d'eux; et un manque d'empressement de la communauté au sens large à l'adoption des PID, que l'on considère (dans la mesure où ils sont pris en compte) comme étant opaques, trop technocratiques, ou encore sur lesquels la communauté projette ses préoccupations et sa méfiance à l'égard des politiques ou des autorités.

Pour relever ces défis, les avantages de l'adoption et de l'intégration des PID doivent être à la fois tangibles et communiqués efficacement. La partie 3.1 fournit des détails sur les priorités de la

collectivité en matière d'intégration des PID. Un leadership s'avère à la fois nécessaire et souhaité pour établir un consensus et garder le cap sur l'objectif à atteindre: une large coalition est proposée pour coordonner la conception et la mise en œuvre de la stratégie sur les PID, qui verra à évaluer les options et à leur adéquation avec les objectifs plus larges; ainsi qu'une organisation faisant office de figure de proue pour réaliser ce travail, possédant une influence politique et culturelle suffisante pour aider à mettre en œuvre la stratégie, en particulier au niveau fédéral.

Il faut accélérer l'adoption des PID en mettant en place une politique et une structure de soutien cohérentes, permettant à la fois de concilier des mesures incitatives concurrentes, tout en s'assurant que l'aide et les avantages de l'adoption des PID se traduisent de façon concrète et bénéficient à l'ensemble de la communauté.

Ce rapport recommande au CCCPID de :

1. Créer une vaste coalition, en élargissant sa portée et son mandat afin d'inclure un plus grand nombre de parties prenantes (par exemple, maximiser la participation des groupes professionnels et de diverses institutions de recherche, ou formaliser la représentation des chercheuses et chercheurs);
2. Développer une vision commune de la valeur ajoutée que représentent les PID afin que les groupes participent ou s'engagent à promouvoir leur adoption;
3. Analyser les besoins et les obstacles identifiés par la collectivité des PID;
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;
5. Fournir les ressources financières, techniques et pour la sensibilisation afin de soutenir le développement et la gouvernance d'une stratégie commune pour l'accès aux PID et la réutilisation des données. Les ressources techniques et de sensibilisation sont indissociables, pensons notamment à la résilience, à la confiance, etc.;
6. Identifier les mesures à prendre à court, moyen et long terme pour concrétiser la vision, et susciter l'engagement de tous les acteurs à leur réalisation;
7. Porter la stratégie sur les PID au Canada et consulter, communiquer, puis consulter à nouveau les acteurs pour s'assurer de l'adhésion à la stratégie.

Plus concrètement, la prochaine étape consiste à établir une feuille de route claire pour la création d'une stratégie nationale sur les PID, notamment :

- Établir, de concert avec un large éventail de parties prenantes, une vision de ce à quoi ressemblera le paysage canadien de l'information sur la recherche dans 5 à 10 ans. Cette vision constituera le cœur de la feuille de route;
- Dans le cadre de cette vision, prioriser un ensemble de PID et documenter les contextes d'utilisation de ceux-ci;
- Travailler à rebours à partir de la vision pour établir les conditions préalables à la réussite (peut-être par le biais d'un exercice « pré-mortem » ou d'autres formes de planification de scénarios pour définir les risques ou les facteurs de confusion);
- Établir une priorisation des ressources pour faire progresser l'adoption et la diffusion des PID de base et les métadonnées associées;

MORE+BRAINS

- Élargir le cadre de gouvernance et démontrer que les PID représentent des solutions aux enjeux de recherche afin de garantir un soutien politique et financier à la mise en œuvre de la stratégie

Table des matières

1 Contexte	17
2 Méthodologie	18
3 Discussion	20
3.1 Conclusions générales	20
3.2 Réflexions	22
4 Recommandations	24

1 Contexte

Ce rapport présente en détail un projet visant à définir les étapes nécessaires à la création d'une stratégie sur les identifiants pérennes (PID) pour la communauté canadienne de recherche. Ce rapport s'appuie sur :

- des années de promotion des PID au Canada;
- la création de deux grands consortiums de PID qui offrent un accès aux PID, DataCite et d'ORCID, ainsi qu'un soutien technique non seulement pour les organismes, mais aussi pour les chercheuses et chercheurs;
- la création du Comité consultatif canadien sur les identifiants pérennes (CCCPID).

Il met en évidence les possibilités et les défis qui devraient faire l'objet d'activités futures, et propose des recommandations pratiques sur la façon dont celles-ci pourraient être priorisées et abordées par la « vaste coalition » élargie des parties prenantes de la recherche au Canada.

En réponse à la demande de soumission du CCCPID pour un consultant chargé d'élaborer une stratégie nationale de PID pour le Canada, MoreBrains Cooperative a présenté avec succès une proposition qui prévoyait un processus en plusieurs étapes pour créer une telle stratégie : une feuille de route pour une stratégie nationale sur les PID au Canada.

Le projet qui en résulte repose sur le principe que les progrès de l'infrastructure de PID du Canada dépendront de l'atteinte d'un consensus sur les normes, les flux de travail, les meilleures pratiques et les approches de gouvernance les plus appropriées parmi les divers intervenants de la recherche au Canada. Chaque milieu de recherche national est unique, car il est façonné par l'histoire et les normes culturelles, ainsi que par les incitatifs actuels et les priorités de la stratégie de recherche. L'établissement d'une véritable adhésion et d'un engagement envers les principes qui sous-tendent la stratégie sur les PID exigera du temps et des consultations, et la stratégie elle-même doit non seulement faire l'objet d'une analyse minutieuse, mais elle doit aussi être élaborée conjointement avec les bailleurs de fonds, les établissements et les éditeurs canadiens, et ce, dans une approche de coconstruction et de sensibilisation.

Depuis novembre 2021, l'équipe de MoreBrains, appuyée par les membres du CCCPID, a travaillé avec un groupe diversifié de parties prenantes afin de définir les exigences et la portée d'une stratégie canadienne à la fois solide et pragmatique en matière de PID. Ce rapport présente les résultats de notre consultation auprès de la communauté, qui avait pour objectifs de :

- déterminer et décrire les objectifs stratégiques de recherche;
- identifier de possibles PID prioritaires, indispensables pour atteindre ces objectifs, avec des recommandations pour toute autre étude nécessaire afin d'évaluer les niveaux d'adoption et de sensibilisation pour ces PID;
- recommander des flux de travail et les meilleures pratiques pour assurer une interconnexion adéquate des PID par le biais des métadonnées;
- établir un dialogue avec la communauté dans son ensemble pour susciter son adhésion à une stratégie sur les PID;
- recenser les besoins non comblés d'ordre technique ou relatifs à la gouvernance;

- identifier les organisations ou groupes de parties prenantes existants susceptibles d'assumer ces responsabilités et/ou toute nouvelle structure ou organisation pertinente;
- mettre en valeur les stratégies de gouvernance, de participation et de sensibilisation auprès de toute structure ou organisation, nouvelle ou existante.

2 Méthodologie

L'équipe de MoreBrains a d'abord entrepris une recherche documentaire, en explorant une série de documents et de ressources préparés par le CCCPID, les consortiums canadiens de PID existants et les diverses initiatives qui y sont liées, ainsi que de la documentation portant sur les stratégies et les services de PID à l'échelle internationale. Cette recherche a été suivie d'une série d'interviews semi-dirigées, dont les résultats ont été utilisés pour orienter la conception de trois ateliers destinés aux parties prenantes.

Cette démarche nous a permis d'explorer le paysage de manière plus large, puis de nous concentrer sur des considérations et des facteurs particuliers à étudier.

Les personnes interviewées ont été choisies par un groupe de dirigeantes et dirigeants du consortiums canadiens ORCID et DataCite, du Réseau canadien de documentation pour la recherche (RCDR) et de l'équipe de GDR de l'Alliance de recherche numérique du Canada (l'Alliance). Les interviews ont été réalisées par Josh Brown, cofondateur de MoreBrains et responsable de la recherche et de la stratégie. Les six personnes interrogées représentaient un éventail de points de vue, issus des grandes catégories suivantes :

- Comités/établissements/membres
- Politique
- Bailleurs de fonds
- Chercheuses et chercheurs

Certaines personnes interviewées ont été ou sont impliquées dans le CCCPID ou dans l'un des groupes de gouvernance des consortiums canadiens de PID existants, alors que d'autres sont moins directement engagées en matière de PID.

Les résultats des interviews ont servi de base à une série de trois ateliers offerts « sur invitation » à des parties prenantes, qui ont eu lieu en mode virtuel entre le 28 février et le 15 mars 2022. Ces ateliers ont permis d'explorer les questions soulevées dans les interviews avec de plus grands groupes de parties prenantes issus de la communauté de recherche canadienne. Les participantes et participants aux ateliers ont été identifiés et sélectionnés par les commanditaires du projet, avec les commentaires et les recommandations de l'équipe de MoreBrains.

On comptait 40 participants provenant de l'ensemble de la communauté de recherche canadienne, issus d'un éventail de groupes de parties prenantes, notamment bailleurs de fonds provinciaux/nationaux, gouvernements, universités (dont des bibliothèques, bureaux de recherche, TI), soins de santé, édition et chercheuses et chercheurs.

Les six personnes interviewées au départ ont été invitées à participer à la série d'ateliers, et ont été réparties sur les trois séances. Le troisième atelier a réuni certains délégués des ateliers 1 et 2 ainsi

que des nouveaux venus, afin de garantir un bon éventail de perspectives sur les questions abordées.

Les commanditaires du projet ont assisté à tous les ateliers – principalement à titre d’observateurs, même si certaines personnes ayant une expertise particulière (stratégie, mise en œuvre) ont participé formellement à la séance correspondante.

Les ateliers se sont déroulés sur Zoom et ont été appuyés par du matériel, des présentations, etc., fournis en anglais et en français. Des employés du RCDR étaient présents pour assurer au besoin la traduction simultanée.

Les discussions en plénière ont été animées et structurées à l’aide de sondages en ligne sur Mentimeter.

Quels points névralgiques constatez-vous actuellement dans le secteur en matière de partage, de réutilisation ou d’échange d’information?



Figure 1. Exemple d’un sondage Mentimeter réalisé dans le cadre des ateliers (NB : cette capture d’écran ne montre que l’anglais)

Les discussions en petits groupes ont permis d’approfondir des questions précises dans un environnement de tableau blanc virtuel, à l’aide de Miro.

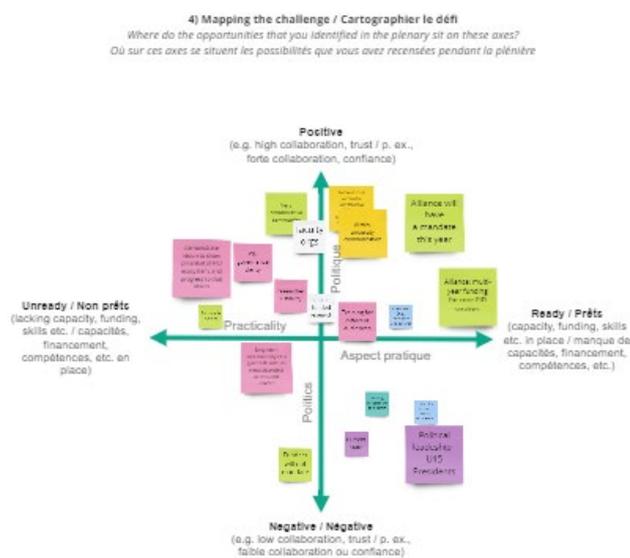


Figure 2. Exemples de documents de discussion issus des séances sur tableau blanc virtuel

Ces discussions, complétées par les recherches documentaires et l'expérience de l'équipe de MoreBrains dans la conduite de projets similaires dans d'autres pays, ont servi à formuler nos recommandations au CCCPID sur les étapes ultérieures à suivre en vue de l'élaboration et du lancement de la stratégie canadienne de mise en œuvre des PID.

3 Discussion

3.1 Conclusions générales

Cette section présente les grandes leçons et les principaux résultats de la phase de recherche du projet.

Certains de ces problèmes sont atténués, dans une certaine mesure, par les principaux atouts identifiés comme étant déjà en place pour soutenir l'évolution de la stratégie canadienne sur les PID :

- On constate beaucoup de dynamisme et d'engagement au sein d'une communauté centrale de PID;
- La communauté de recherche canadienne bénéficie d'une expertise sur les questions pratiques et politiques relatives aux PID, et entretient d'excellents rapports avec d'autres consortiums, expertes et experts et fournisseurs de PID ailleurs dans le monde;
- Le Canada a l'occasion de jouer un rôle de premier plan dans l'équilibre entre les données ouvertes et la souveraineté des données (principes FAIR et CARE). Bien qu'innover comporte certains risques, l'équilibre risque/récompense sera favorable si la stratégie sur les PID est inclusive dans sa conception et renforcée par un soutien opérationnel et financier, afin de favoriser la pleine participation de toutes les communautés.

Notre analyse des trois ateliers en particulier a fait ressortir un certain nombre de facteurs pertinents pour les premières étapes de l'élaboration d'une stratégie nationale sur les PID pour le Canada :

- Les groupes de parties prenantes présentent des niveaux de confiance différents. Cela pourrait limiter l'adhésion au début, et il faudra assurer une consultation et une communication constantes pour y remédier;
- L'infrastructure est fortement fragmentée entre les organismes de financement. Il est donc difficile pour les chercheuses et chercheurs de comprendre ce que l'on attend d'eux, cela augmente les coûts pour les fournisseurs de services et il est plus difficile de faire une analyse de rentabilité pour l'adoption ou l'utilisation d'un PID en particulier;
- On observe un manque d'empressement de la communauté au sens large à l'adoption des PID, que l'on considère (dans la mesure où ils sont pris en compte) comme étant opaques, trop technocratiques, ou encore sur lesquels la communauté projette ses préoccupations et sa méfiance à l'égard des politiques ou des autorités;

- La compréhension et l'acceptation d'ORCID et d'autres PID ne sont pas uniformes au sein des communautés et entre elles (voir les points précédents), et on ne pourra finalement y remédier que par l'éducation et des messages cohérents et soigneusement conçus.

Les ateliers et les interviews ont tous deux fait ressortir un ensemble commun de perspectives fondamentales.

Les avantages doivent être à la fois concrets et communiqués de façon efficace.

Les bailleurs de fonds et les institutions de recherche doivent s'assurer que les chercheuses et chercheurs voient d'abord des avantages, et non de nouvelles contraintes inhérentes à l'implantation des PID.

Puisque la transition pourrait entraîner un nouveau fardeau administratif ou du moins une gestion du changement importante, par exemple, l'apprentissage d'une nouvelle façon d'entreprendre une tâche comme la saisie complète d'un enregistrement ORCID, il peut être plus réaliste de dire que les avantages doivent être perçus comme étant supérieurs aux inconvénients.

Les priorités qui ont été dégagées lors des ateliers peuvent aider à établir les voies à suivre pour démontrer ces avantages, notamment :

- l'intégration dans les systèmes institutionnels;
- des systèmes de financement avec des intégrations de PID améliorées et (surtout) cohérentes;
 - flux bidirectionnel vers ORCID, etc.;
 - enregistrement des ID de subventions;
- des solutions de partage de données spécifiques au Canada (c'est-à-dire un stockage de données basé au Canada) afin de simplifier les liens avec les registres de PID et d'augmenter le partage de données entre les institutions de recherche et les bailleurs de fonds.

Un leadership est à la fois nécessaire et souhaité.

Concrètement, à cette étape, le « leadership » consiste à diriger un processus basé sur le consensus et à maintenir l'orientation du programme. Il est manifestement nécessaire et souhaitable qu'une large coalition coordonne la conception et la mise en œuvre de la stratégie sur les PID, en évaluant les options et en réfléchissant à l'adéquation d'interventions précises à des objectifs plus larges. Il faudrait également un organisme disposant d'un capital politique et culturel suffisant pour aider à la réalisation de ce programme, notamment au niveau fédéral, qui puisse servir de figure de proue pour ce travail.

Les discussions tenues lors des ateliers sur les obligations émises par les bailleurs de fonds en matière de PID constituent un exemple intéressant. Historiquement, l'exigence de fournir un PID a permis d'augmenter la portée de celui-ci, comme dans le cas de l'obligation de fournir un identifiant ORCID imposée par le Wellcome Trust aux demandeurs de subventions. Mais ce caractère obligatoire a également eu des conséquences inattendues; par exemple, bon nombre d'enregistrements ORCID sont désormais vides, ayant été créés en grande partie pour se conformer à une obligation, puis abandonnés. Les questions à considérer concernant l'obligation de fournir un PID sont les suivantes : S'agit-il d'une approche unique, qui convient à tous ? Qu'est-ce qui

fonctionnerait le mieux, l'adhésion ou l'obligation ? Comme les obligations peuvent prendre de nombreuses formes (exiger des bénéficiaires de subventions qu'ils citent un DOI pour la subvention dans les remerciements de financement, ou encore, exiger que les dépôts enregistrent des DOI pour tous les travaux d'un type particulier), il faudrait que la communauté adhère de façon consensuelle aux principes de collecte de données institués dans la stratégie avec une évaluation.

La première tâche à accomplir, identifiée lors des ateliers, est d'établir un leadership qui dirigera ce processus, avec la plus large consultation possible. Une fois celui-ci en place, la priorité devrait être de parvenir à un accord sur une stratégie avant de décider de tactiques précises, telles que des exigences pour les PID.

3.2 Réflexions

Au cours du processus de recherche, nous avons dégagé trois thèmes importants qui ont une incidence majeure sur les prochaines étapes de l'élaboration d'une stratégie nationale sur les PID pour la communauté canadienne de la recherche.

L'adoption

Pour que les PID et leurs métadonnées associées soient utiles et fiables, ils doivent être présents et disponibles pour la majorité des entités qu'ils servent à identifier. La couverture de tout système de PID donné est déterminante pour son utilité. Les PID doivent être enregistrés et partagés le plus tôt possible dans le cycle de vie d'une entité, ce qui signifie que les outils utilisés pour enregistrer ou créer ces entités doivent disposer d'intégrations de PID parfaitement fonctionnelles. En même temps, les métadonnées qui leur sont associées doivent être accessibles sous des formes structurées et réutilisables pour être intégrées dans des systèmes administratif, de découverte, de rapport et d'analyse.

À l'heure actuelle, les avantages des PID sont essentiellement abstraits et le resteront jusqu'à ce que les PID et leurs métadonnées associées soient utilisés et réutilisés de manière récurrente et cohérente dans les systèmes, les plateformes et les outils du réseau de recherche canadien.

Les recherches que nous avons menées pour ce projet ont montré qu'il n'existe tout simplement pas assez d'intégrations techniques des PID à l'heure actuelle. En conséquence, les chercheuses et chercheurs doivent ajouter manuellement les mêmes informations dans plusieurs systèmes de financement et/ou d'institutions de recherche, et ce, à plusieurs reprises.

Selon des participants au projet, le coût et la complexité de l'intégration des PID constituent un obstacle majeur à l'adoption. De nombreuses petites institutions de recherche ne peuvent pas se permettre l'infrastructure technique nécessaire, et les implantations uniques « faites maison » entraînent une duplication des efforts et, par conséquent, un coût plus élevé que ce que le secteur peut supporter. En même temps, les participants ont signalé que de nombreux systèmes commerciaux qui offrent des intégrations de PID (par exemple, CRIS) sont trop coûteux pour un certain nombre d'institutions de recherche.

L'une des suggestions qui sont ressorties des différentes discussions en ateliers est la création d'une solution canadienne de partage de l'information, hébergée localement, afin de permettre l'accès aux API et aux métadonnées des PID, combinées à d'autres informations. Cela pourrait prendre la forme

d'une banque de métadonnées pouvant se connecter aux API de PID, aux systèmes des bailleurs de fonds et aux institutions. Cette idée a été suggérée à l'origine lors du deuxième atelier (prolongeant une discussion semblable en cours depuis un certain temps³), puis approuvée lors du troisième atelier. Cela pourrait constituer un moyen intéressant de réduire les coûts, de favoriser l'adoption et de garantir un accès équitable aux API et aux services de PID.

La cohésion

L'absence de démarches cohérentes et constantes en matière de PID au sein des groupes de parties prenantes et entre eux constitue une lacune importante qui doit être comblée, et c'est pourquoi le CCCPID a commandé cette étude. Bien que, comme mentionné ci-dessus, il existe un noyau de personnes très mobilisées et engagées dans cette initiative, la portée de ce groupe doit être renforcée pour que les prochaines étapes du processus d'« optimisation des PID » soient réalisées de manière efficace. À l'heure actuelle, l'adoption des PID est freinée par l'absence d'une vision et d'objectifs nationaux unifiés, comme en témoignent les variations dans les politiques d'un bailleur de fonds à un autre, les incitatifs incohérents et le manque de priorités politiques claires. Ces incohérences entre secteurs sont devenues des obstacles à la mise en place d'un environnement de recherche efficace et optimisé par les PID.

Il s'avère donc nécessaire de reconfigurer et d'harmoniser les priorités et les activités afin d'assurer la réussite de la stratégie sur les PID. Pour y parvenir, il faudra s'attaquer aux problèmes historiques et favoriser un engagement constructif à l'égard de la capacité potentielle des PID à répondre aux sources de frustration de la communauté des chercheuses et chercheurs. Un exemple qui est revenu souvent dans les interviews et les ateliers est celui du CV commun canadien. Un certain nombre de participants nous ont indiqué que le CVC crée une charge administrative, et il a manifestement laissé de profondes cicatrices chez nombre de ses utilisateurs.

Ces problèmes peuvent empêcher certains groupes de parties prenantes de s'engager dans le processus de la stratégie sur les PID ou d'agir en temps opportun. Au cours des discussions, certains bailleurs de fonds ont exprimé des inquiétudes quant à l'avancement des solutions en raison des risques pour la réputation. Un processus de priorisation cohésif, ouvert et inclusif contribuera à réaliser l'analyse de rentabilité d'un tel projet, en veillant à ce que les priorités soient partagées et soutenues par une multitude de parties prenantes et en aidant à apaiser leurs inquiétudes.

L'équité

Une des leçons tirées d'autres initiatives nationales sur les PID est qu'aucun groupe, organisation ou communauté ne tirera pleinement profit des PID tant que tous n'en bénéficieront pas⁴. Les effets de réseau et les économies d'échelle qui sont à l'origine de ces avantages reposent sur une participation la plus complète possible.

³ Barsky, E. (14 décembre 2020). Research Information Management Integration for Canada [R]. doi: <http://dx.doi.org/10.14288/1.039527> « Nous proposons que la NOIRN joue un rôle de premier plan dans la mise en œuvre d'une intégration de la gestion des informations relatives à la recherche à l'échelle nationale, légère et interopérable pour le Canada, en tant que composante essentielle de l'écosystème global de la GDR. »

⁴ Voir, par exemple : Josh Brown, Phill Jones, Alice Meadows, Fiona Murphy et Paul Clayton. (2021). UK PID Consortium: Cost-Benefit Analysis (1.0). Zenodo. <https://doi.org/10.5281/zenodo.4772627>

Tout cela comporte de nombreuses dimensions : petites et grandes institutions, échelles locale, provinciale et nationale, communautés anglophones et francophones, bourses autochtones, différences entre les disciplines, etc. D'autres pays qui examinent des stratégies sur les PID rencontrent des difficultés similaires, aussi existe-t-il des points communs et une expérience sur lesquels nous pouvons nous appuyer. Le Canada présente un mélange unique de différences culturelles et des langues multiples; si certains aspects sont semblables à d'autres pays, comme l'Australie, il se distingue sans contredit par son histoire et son contexte.

Les discussions menées lors des interviews et des ateliers ont fait ressortir certains des défis que représente la langue à plusieurs niveaux. La langue peut dénoter des groupes culturels, elle peut aussi élever des obstacles à l'accès par le biais d'interfaces unilingues, et elle peut également compliquer la réutilisation des métadonnées ou des informations.

D'autres défis culturels et politiques doivent être abordés directement, notamment les risques de paternalisme, d'appropriation ou de sous-représentation des peuples autochtones, ainsi que les préoccupations relatives à la vie privée et à la quantification (proposées comme hypothèse pour expliquer la lenteur de l'adoption des identifiants ORCID, notamment au sein des communautés francophones).

Il sera essentiel de jeter des ponts entre les groupes, les cultures et les communautés si l'on veut que la stratégie sur les PID soit couronnée de succès et que tous ceux et celles qui participent ou contribuent à la recherche canadienne puissent en bénéficier. Même si les défis sont spécifiques au Canada, nous avons l'occasion de montrer l'exemple à l'échelle mondiale – cela exigera toutefois un modèle de souveraineté des données élaboré de manière transparente et inclusive.

4 Recommandations

En conclusion, nous recommandons au CCCPID de :

1. Créer une vaste coalition, en élargissant sa portée et son mandat afin d'inclure un plus grand nombre de parties prenantes (par exemple, maximiser la participation des groupes professionnels et de diverses institutions de recherche, ou formaliser la représentation des chercheuses et chercheurs);
2. Développer une vision commune de la valeur ajoutée que représentent les PID afin que les groupes participent ou s'engagent à promouvoir leur adoption;
3. Analyser les besoins et les obstacles identifiés par la collectivité des PID;
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;
5. Fournir les ressources financières, techniques et pour la sensibilisation afin de soutenir le développement et la gouvernance d'une stratégie commune pour l'accès aux PID et la réutilisation des données. Les ressources techniques et de sensibilisation sont indissociables, pensons notamment à la résilience, à la confiance, etc.;
6. Identifier les mesures à prendre à court, moyen et long terme pour concrétiser la vision, et susciter l'engagement de tous les acteurs afin de les réaliser;
7. Porter la stratégie sur les PID au Canada et consulter, communiquer, puis consulter à nouveau les acteurs pour s'assurer de l'adhésion à la stratégie.

Plus concrètement, la prochaine étape consiste à établir une feuille de route claire pour la création d'une stratégie nationale sur les PID, notamment :

- Établir, de concert avec un large éventail de parties prenantes, une vision de ce à quoi ressemblera le paysage canadien de l'information sur la recherche dans 5 à 10 ans. Cette vision constituera le cœur de la feuille de route;
- Dans le cadre de cette vision, prioriser un ensemble de PID et documenter les contextes d'utilisation de ceux-ci;
- Travailler à rebours à partir de la vision pour établir les conditions préalables à la réussite (peut-être par le biais d'un exercice « pré-mortem » ou d'autres formes de planification de scénarios pour définir les risques ou les facteurs de confusion);
- Établir une priorisation des ressources pour faire progresser l'adoption et la diffusion des PID de base (et des métadonnées associées);
- Élargir la base de données probantes et le cadre de gouvernance afin de garantir le soutien politique et financier qui permettra la mise en œuvre de la stratégie.