

Les données et l'innovation au Canada

Contexte pour une dialogue nationale

Élaboré par le comité directeur de Données de recherche Canada

1er mai 2017 | Version 3.0

DOI: 10.5281/zenodo.1135183

rdc-drc.ca | @rdc-drc

La recherche-développement est un des grands moteurs de l'économie canadienne. Ainsi en va-t-il un peu partout dans le monde, et les données en sont indissociables. Données de recherche Canada (DRC) collabore avec de nombreuses parties prenantes pour faire en sorte que les Canadiens aient accès aux données issues des recherches financées par les deniers publics et conservées dans les collections nationales ou internationales, afin qu'elles alimentent l'innovation dont chacun retire les fruits. Pour optimiser les retombées socioéconomiques de la recherche publique, DRC s'efforce de mieux harmoniser les activités des intervenants, de promouvoir une meilleure réutilisation des données et de combler les lacunes des technologies habilitantes. Ces activités s'inscrivent dans les efforts déployés à l'échelon national pour faire progresser l'infrastructure numérique de la recherche au pays.

Le document que voici a été rédigé au nom des parties prenantes de DRC, pour faciliter le développement de services qui aideront les chercheurs à exploiter les données scientifiques tout au long de leurs travaux.

Le document « *Les données et l'innovation au Canada* » remet en contexte le dialogue national que poursuivent les intervenants de l'écosystème canadien de la recherche. Ce dialogue contribuera à échafauder la stratégie qui engendrera une infrastructure nationale et des services tenant compte des intérêts et des capacités des provinces et des territoires. DRC focalisera la vision nationale en organisant des consultations et des discussions qui la préciseront au cours des mois à venir et permettront aux parties prenantes de s'aligner avec un jeu commun de principes et de fonctions élémentaires autour duquel se cristallisera le cadre national en gestion des données. DRC part de l'hypothèse que les technologies numériques ont eu un effet prononcé sur la méthodologie scientifique et pourraient vraiment transformer la manière dont sont effectuées des recherches, donc la façon dont les nouvelles connaissances font progresser la société et la place que chacun y occupe. DRC arrime cette vision à un ensemble de facteurs sur lesquels il faut se pencher si l'on veut épauler les chercheurs dans leurs tentatives pour découvrir, analyser, diffuser et réutiliser les données dans tous les domaines.

UTILITÉ DES DONNÉES, DE LEUR CONSULTATION À L'OBTENTION DE RÉSULTATS

- **Ensemble des intervenants** : accès immédiat à des données bien définies et étayées; bonification ou validation des recherches en cours; enrichissement du savoir
- **Chercheurs** : découverte de collaborations inconnues auparavant; découverte de données utiles jusqu'à présent dissimulées
- **Organismes de financement** : constitution du profil d'examineurs éventuels; optimisation des sommes injectées dans la recherche active
- **Décideurs** : identification de l'impact des fonds destinés à la recherche thématique ou générale; documentation des politiques; prise de décisions s'appuyant sur des données factuelles

Réutilisabilité

DRC collaborera avec les organisations locales, provinciales, régionales, nationales et internationales afin de rendre les données plus accessibles et d'en accroître la réutilisation. La réutilisabilité demeure l'objectif ultime de services de gestion des données scientifiques robustes et durables qui alimenteront un écosystème de la recherche efficace. Dans cette optique, nous faciliterons la réutilisation des données grâce à des politiques d'encadrement qui concourront à l'adoption d'une approche d'ouverture par défaut pour la dissémination des résultats. De cette façon, on alimentera l'innovation davantage et l'expansion de l'économie s'en trouvera accélérée. Investir dans les services et les ressources qui appuieront pareils efforts au niveau du financement de la gestion des données scientifiques et de l'infrastructure numérique de la recherche en fait partie intégrante. On suggère souvent de réserver 1,5 à 2 % des fonds totaux destinés à la recherche au soutien de tels efforts. Selon une étude récente¹, les données ouvertes ajoutent annuellement 0,5 % de plus au PIB que les données protégées ou consultables contre rémunération. L'investissement se paierait donc de lui-même.

Découverte

L'écosystème de la recherche canadien doit impérativement s'appuyer sur la découverte, qui est la clé pour atteindre l'objectif de réutilisabilité. Les systèmes de découverte actuels ne fonctionnent souvent qu'au niveau de l'institution ou, dans certains cas, entre plusieurs institutions, voire au palier provincial ou régional. Il faut en garantir la fonctionnalité à la grandeur du pays. Il faut également concevoir une couche nationale pour la découverte, c'est-à-dire un système qui donnera accès à tous les résultats de la recherche par une interface harmonieuse, afin qu'on puisse répondre aux questions posées par n'importe quel utilisateur, qu'il s'agisse d'un fonctionnaire, du cadre d'une université, d'un chercheur ou d'un simple curieux. Les Canadiens devraient pouvoir accéder aisément aux données scientifiques pour trouver la réponse à leurs interrogations. La capacité de découvrir les preuves scientifiques est fondamentale à l'avancement du savoir et à l'innovation, donc à faire progresser la société canadienne.

Métadonnées

Une approche vraiment révolutionnaire à la découverte devrait englober tous les résultats scientifiques pendant le cycle entier de la recherche. Les résultats seraient connectés par des éléments ou des acteurs communs (chercheurs-boursiers, ressources scientifiques, services, bailleurs de fonds et produits classiques comme les articles de périodiques et les données). En décrivant et en recueillant le profil complet de chaque activité de recherche, on bâtira une base de métadonnées qui exploitera la puissance des nouvelles technologies informatiques et fera ressortir les détails les plus infimes des recherches financées par l'État. On minimisera les difficultés éventuelles si cette base de métadonnées s'appuie sur des normes communes et des systèmes interopérables. Un aspect capital d'un système stockant des résultats aussi variés serait de veiller à ce que les moyens adoptés pour garantir la sécurité des métadonnées et des données qui s'y associent se reflètent dans les fonctions qui commandent l'accès aux données et les particularités du système.

¹ <http://theodi.org/research-economic-value-open-paid-data>

Préservation

Pour que les données soient préservées, on doit s'assurer que les résultats des recherches individuelles seront conservés longtemps, le but ultime de la découverte étant qu'on puisse consulter les données et les informations factuelles. Une couche appropriée de stockage et de préservation, qu'on pourrait appeler « fiducie nationale de données », serait la base essentielle à une gestion efficace des données scientifiques. Comme c'est le cas pour les collections analogiques que les bibliothèques nationales et institutionnelles préservent depuis des siècles, et dont elles assument l'intendance, il faut reconnaître la nécessité d'investir dans une infrastructure à long terme analogue pour les actifs numériques.

DRC estime que l'accès aux résultats des recherches est aussi crucial que les réseaux ultrarapides si le Canada veut rester à la fine pointe de la recherche et de l'innovation. Nous avons pris l'habitude de soutenir financièrement et de manière prolongée le développement et le maintien des réseaux; la même attitude doit s'appliquer aux infrastructures de stockage. Si la préservation des résultats scientifiques ne suffit pas à justifier une intendance durable de la documentation, la capacité d'exploiter les données pour alimenter la recherche, la découverte et la création de politiques le fera certainement.

Coordination

Diverses organisations dialoguent sur le développement de services nationaux de données. Grâce à DRC, les agences de recherche, les universités et les administrations publiques peuvent se regrouper pour mieux se coordonner et échauffer la stratégie qui mettra en place les éléments de l'infrastructure de recherche numérique canadienne se rapportant aux données, ainsi que pour établir les sommes qui serviront à cette fin. Cet effort de coordination doit aussi porter sur l'élaboration d'une stratégie de longue haleine montrant clairement que la seule bonne manière de procéder consiste à investir à long terme, pas au cas par cas.

Enfin, puisque la majorité des fonds destinés à la recherche émanent de sources fédérales et provinciales et puisque la récente Politique des trois organismes sur le libre accès aux publications aborde la question de la gestion des données, DRC est conscient que les chercheurs redoutent un alourdissement du fardeau administratif et l'incapacité, pour leurs institutions, de satisfaire aux nouvelles exigences. Certes, les avantages surpassent les difficultés éventuelles, cependant, sans vraie gouvernance et si l'on ne tient pas correctement compte de ce que de tels changements signifient pour les chercheurs et leurs institutions, les changements transformationnels de cette nature ne se concrétiseront tout simplement pas. Il est primordial d'établir où le financement peut accélérer le mouvement et d'harmoniser les activités du Canada avec celles de ses partenaires de l'étranger. Une approche coordonnée et durable à la transformation des réserves nationales de données est indispensable si on veut favoriser la découverte et la réutilisation des données scientifiques et des métadonnées interopérables bien préservées. DRC s'est engagé à faciliter cette transformation et à œuvrer avec ses partenaires du Canada et de l'étranger afin d'atteindre ce but.