

LA GESTION DES DONNÉES DE RECHERCHE DANS LES UNIVERSITÉS CANADIENNES

DÉCLARATION DE PRINCIPES

MARS 2016

Contexte

La déclaration de principes sur la gestion des données de recherche (GDR) que voici a été élaborée par un groupe de travail composé de cadres supérieurs administratifs des universités et de membres de Données de Recherche Canada (DRC). Elle fait suite à une rencontre parrainée par DRC et l'Université de l'Alberta, tenue à Edmonton le 17 novembre 2015 grâce à une aide financière dispensée par le CRSH, au nom des organismes subventionnaires canadiens. Y assistaient des cadres supérieurs de seize institutions et les représentants des conseils subventionnaires fédéraux ainsi que des organisations internationales s'intéressant aux données de recherche. Les personnes qui se sont rencontrées à cette occasion étaient toutes au fait des problèmes liés à la GDR, en partie à cause des efforts actuellement déployés par les organismes subventionnaires publics pour élaborer des lignes directrices dans ce domaine. Voir <http://www.science.gc.ca/default.asp?lang=Fr&n=83F7624E-1>.

Les personnes réunies à Edmonton ont confié la rédaction d'une déclaration de principes à un groupe de travail¹ qui a commencé par examiner plusieurs énoncés internationaux pertinents avant de retenir les principes formulés par les Research Councils of the United Kingdom (www.rcuk.ac.uk/research/datapolicy) comme point de départ. Les auteurs du présent document remercient les RCUK pour cet apport.

Chaque université élaborera une politique sur la GDR qui répondra à ses besoins et à ses contraintes. Cependant, il est possible de s'inspirer de ce qui a déjà été fait et de collaborer afin de ne pas avoir à réinventer la roue à chaque fois. En adoptant collectivement les principes fondamentaux énoncés plus

¹ Le groupe de travail se composait comme suit : Mark Leggott, bibliothécaire, Université de l'Île-du-Prince-Édouard (directeur exécutif désigné de Données de recherche Canada); Frédéric Bouchard, vice-recteur associé, Université de Montréal; Robert Haché, vice-président à la recherche et à l'innovation, Université York; Gerald Beasley, vice-recteur et bibliothécaire en chef, Université de l'Alberta; Susan Babcock, directrice exécutive par intérim du Bureau de la déontologie scientifique, Université de l'Alberta; David Castle, vice-président à la recherche, Université de Victoria; Andrew Bjerring, président du Comité des politiques de DRC; Walter Stewart, coordonnateur de DRC.

loin, on facilitera considérablement l'élaboration de nouveaux outils, services et infrastructures, tout en se rapprochant des objectifs avoués d'une science plus ouverte et d'un partage plus démocratique des données. Quelques universités se sont déjà dotées d'une politique de GDR; d'autres en ont entrepris le développement et d'autres encore ne font qu'y songer. Cette déclaration de principes n'a pas pour but d'instaurer une politique nationale en la matière, ni de remplacer les politiques existantes des universités, mais bien de proposer un ensemble commun de principes gouvernant la GDR dans les universités du Canada.

DÉCLARATION DE PRINCIPES SUR LA GESTION DES DONNÉES DE RECHERCHE DANS LES UNIVERSITÉS CANADIENNES

Préambule

Parce qu'elles se sont engagées à atteindre l'excellence en recherche et à partager le savoir, les universités canadiennes sont déterminées à surmonter les obstacles qui empêchent une gestion efficace des données de recherche ainsi qu'à saisir les occasions encourageant la diffusion des connaissances. Les principes énoncés ci-dessous épousent le glossaire développé par Données de recherche Canada, qu'on pourra consulter sur <http://www.rdc-drc.ca/fr/glossaire>.

Plus précisément, l'expression « données de recherche » a le sens que voici :

Données essentiellement employées pour alimenter les enquêtes techniques ou scientifiques, la recherche, les études ou des activités artistiques. On s'en sert pour étayer la méthode expérimentale ou les chercheurs estiment communément qu'elles sont nécessaires pour valider leurs observations et les résultats de leurs travaux. Tout contenu numérique ou pas est susceptible de devenir des données de recherche. Les données de recherche peuvent être de nature expérimentale ou opérationnelle, émaner d'observations ou d'une tierce partie, être issues du secteur public, venir de la surveillance et comprendre des données traitées ou recyclées.

Le terme « métadonnées » désigne ce qui suit :

Au sens propre, « données sur les données ». Données qui définissent et décrivent les particularités d'autres données. On s'en sert pour faciliter la compréhension des données techniques et opérationnelles ainsi que des processus se rapportant aux données. Les métadonnées opérationnelles comprennent le nom et la description des domaines, les entités et les attributs, la nature des données d'attribut et d'autres propriétés des attributs, la description des intervalles, les valeurs acceptables pour les domaines et leur définition. Les métadonnées techniques comprennent le titre des tableaux et des colonnes des bases de données physiques, les propriétés des colonnes et les propriétés des autres objets de la base de données, notamment la façon dont les données sont stockées. Les métadonnées sur les processus définissent et décrivent les caractéristiques d'autres éléments du système (processus, règles opérationnelles, programmes, tâches, outils, etc.). Les métadonnées sur la gérance des données portent sur les responsables des données, les méthodes de gestion des données et les responsabilités attribuées à chacun.

Par « gestion des données de recherche » (GDR), on entend :

Stockage, consultation et préservation des données issues d'un travail de recherche quelconque. La gestion des données touche le cycle de vie complet des données, de la planification des recherches à leur exécution et de l'enregistrement des données au moment de leur création et de leur exploitation jusqu'à la préservation à long terme des résultats, une fois les recherches achevées. Au nombre des activités et des difficultés associées à la gestion des données figurent les suivantes : désignation des fichiers (manière adéquate de nommer les fichiers de données); contrôle et assurance de la qualité des données; accès aux données; documentation (y compris degré d'incertitude); création de métadonnées et de vocabulaires contrôlés; stockage des données; archivage et préservation des données; partage et réutilisation des données; intégrité des données; protection des données; protection des renseignements personnels; droits associés aux données; protocoles expérimentaux (en laboratoire ou sur le terrain).

Principes

1. Importance des données pour la recherche

Les universités canadiennes reconnaissent le rôle capital des données pour la recherche au 21^e siècle. Les données sont à la fois le fruit et les fondations de la recherche. Les gérer convenablement et en faciliter l'accès d'une manière appropriée forment la base de l'érudition et de la découverte dans les sciences aujourd'hui.

2. Collaboration au Canada et à l'étranger

Les universités canadiennes s'efforceront collectivement de surmonter les obstacles qui entravent la GDR ainsi qu'à saisir les occasions de développement qui s'y rattachent en s'alignant sur les activités internationales pertinentes, notamment le développement et l'adoption de normes grâce auxquelles les données de recherche canadiennes deviendront interopérables avec celles de partenaires à l'étranger.

3. Accès

Les données issues des recherches financées par les fonds publics sont de ce fait un bien public, produit dans l'intérêt de la collectivité. Elles seront donc normalement accessibles en temps opportun, avec le moins de restrictions possible. En règle générale, les droits sur les données des recherches bénéficiant d'un soutien public ne seront pas cédés à autrui sans la garantie qu'elles resteront accessibles et que l'on pourra s'en resservir.

4. Déontologie, réglementation et protection des renseignements personnels

Les universités canadiennes admettent que des ententes contractuelles pourraient être assorties de considérations relevant de la protection des renseignements personnels, d'ordre juridique ou déontologique et d'intérêts commerciaux susceptibles de limiter la diffusion des données. Elles élaboreront donc des politiques qui tiendront compte de ces enjeux.

5. Usage privilégié

Ceux qui poursuivent les recherches disposeront d'une période durant laquelle ils jouiront de l'usage privilégié des données qu'ils ont récoltées ou produites, par exemple pour leur permettre de diffuser les résultats de leurs travaux. La durée de cette période pourra varier selon la nature de la discipline universitaire concernée.

6. Reconnaissance de la contribution intellectuelle

Des politiques institutionnelles et des mesures de reconnaissance souligneront l'apport intellectuel des chercheurs qui produisent, préservent et partagent des données. Les utilisateurs des données existantes mentionneront la source des données et respecteront les conditions associées à leur exploitation.

7. Intérêt public

Puisque les données de recherche sont un bien public, leur gestion se fera dans l'intérêt de la collectivité. Bien que la GDR incombe d'abord au chercheur principal pendant le projet, la préservation à long terme des données n'est réalisable qu'avec un financement suffisant des organismes subventionnaires et des administrations publiques. Les ressources à ce titre devront être à la hauteur du soutien initial dispensé pour la recherche, notamment sur les plans de la formation et du savoir-faire.

8. Plans de gestion des données

Les plans de gestion des données des institutions et ceux propres à un projet suivent habituellement des normes internationales et des pratiques exemplaires reconnues en la matière. Pareils plans devraient reconnaître la valeur potentielle des données à long terme, parfois à des fins distinctes de celles pour lesquelles elles ont été créées au départ. Ils devraient donc intégrer des dispositions et des ressources pour la préservation et la consultation de ces données. Les décisions concernant la durée de conservation des données devraient s'appuyer sur des politiques qui admettent la valeur potentielle à long terme de telles données.

9. Métadonnées et facilité de découverte

En règle générale, on sauvegardera les métadonnées dans un format reconnu à l'échelle internationale et on y donnera librement accès. Ainsi, les données de recherche pourront être plus facilement découvertes et exploitées efficacement par d'autres. La publication des résultats inclura des précisions sur la façon d'accéder aux données d'origine. Si les données ne peuvent être consultées ou ne sont pas encore disponibles (voir les principes 4 et 5), on pourra publier les métadonnées afin de signaler l'existence des données en question aux utilisateurs éventuels.

10. Accès multilingue

Les chercheuses et chercheurs ainsi que les institutions du Canada utiliseront des outils de gestion des données en français ou en anglais, mais idéalement dans ces deux langues.