

## Commentaires de Portage sur l'ébauche de la Politique des trois organismes sur la gestion des données de recherche

[ÉBAUCHE DE LA POLITIQUE](#) des trois organismes originale et [FAQ](#)

### Contexte

Le Réseau Portage, initiative de l'Association des bibliothèques de recherches du Canada (ABRC), se consacre à l'avancement de la gestion des données de recherche (GDR) au Canada. Grâce à son réseau national de plus de 115 experts provenant de plus de 40 organismes, Portage collabore de manière proactive avec des partenaires communautaires afin d'élaborer des services et une infrastructure nationale de GDR à l'appui de tous les chercheurs au Canada.

Portage a eu le privilège de travailler en étroite collaboration avec les représentants des trois organismes au cours de la phase de consultation préalable à l'élaboration des politiques, et de parcourir le pays en janvier et février 2018 pour rencontrer 14 groupes d'intervenants différents dans quatre villes. Ces réunions ont été l'occasion idéale de partager les exigences de la Politique des trois organismes avec les chercheurs et de décrire les services et l'infrastructure de Portage conçus pour aider les chercheurs à répondre à chaque exigence. De plus, ces réunions ont donné à Portage, ainsi qu'à nos collègues des trois organismes, l'occasion d'entendre les commentaires des groupes d'intervenants et de répondre ensemble aux questions soulevées.

Portage a également profité de l'occasion pour convoquer trois assemblées locales en juillet 2018, auxquelles ont participé plus de 32 personnes. Les commentaires recueillis lors de ces assemblées locales et d'autres discussions moins formelles ont été pris en considération lors de la rédaction de la réponse de Portage fournie ci-dessous. Un résumé des commentaires de la collectivité est fourni à l'annexe A.

### Introduction

Portage apprécie la réflexion et le travail acharné consacrés à la formulation de l'ébauche de la Politique des trois organismes sur la GDR. Portage appuie pleinement les trois organismes dans leurs orientations stratégiques en matière de GDR et se réjouit de l'occasion de formuler des commentaires pendant cette période de consultation publique. Les commentaires sont organisés ci-dessous, en suivant les sections du projet de politique.

### Commentaires spécifiques liés à la Politique :

#### 1. Préambule

Les changements suggérés sont indiqués en **rouge et soulignés** ci-dessous :

« Plan de gestion des données – plan qui aide les chercheurs à déterminer la **logistique**, les coûts, les bénéfices et les défis **associés à** la gestion des données et devrait être consulté et mis à jour tout au long du projet de recherche. »

« Par son [Plan d'action du Canada pour un gouvernement ouvert](#), le gouvernement du Canada vise à maximiser l'accès aux résultats de la recherche **qu'il finance** afin

d'encourager une collaboration et un dialogue accrus avec la communauté scientifique, le secteur privé et le grand public. »

COMMENTAIRES : Le préambule présente et appuie fortement le concept de la GDR et offre des relations clés avec d'autres politiques et orientations stratégiques du gouvernement. Dans le préambule, on dit :

- « appuient l'excellence en recherche »
- soient « aussi accessibles que possible »
- « ... données numériques de la recherche est devenue essentielle... »
- « La gestion des données de recherche est nécessaire »
- L'utilisation d'un tel plan est devenue « une pratique exemplaire internationale » et est « essentielle à l'excellence en recherche »
- « nécessité d'établir des politiques pour favoriser leur excellente gestion (de données) »
- « maximiser l'accès aux résultats de la recherche qu'il finance »
- « adopter les politiques destinées à favoriser l'administration efficace des données »
- « favoriser d'excellentes pratiques de gestion des données »

La collectivité était d'avis que l'ébauche de politique ne respectait pas entièrement ces expressions d'appui à la GDR et que, par conséquent, ces idéaux devraient se refléter dans un ensemble plus ferme et plus direct d'exigences stratégiques.

## 2. Objectif de la politique

Aucun commentaire.

## 3. Énoncé de politique

TEXTE ORIGINAL : « La politique vise les détenteurs de subvention et les établissements qui administrent les fonds des trois organismes subventionnaires. Elle ne s'applique pas aux boursiers ou aux titulaires de chaire. »

COMMENTAIRES : Il y a là une contradiction logique qui ne sera pas négligée par les demandeurs de subventions obligés de suivre la nouvelle politique. Il faut expliquer le fait que l'ébauche de politique exclut les « boursiers ou les titulaires de chaires ». Si l'intention est de faire en sorte que ces bénéficiaires des fonds des trois organismes soient éventuellement intégrés à la politique, il faut le préciser.

TEXTE SUGGÉRÉ : « Cette politique ne s'applique actuellement qu'aux détenteurs de subventions et aux établissements qui administrent les fonds des trois organismes subventionnaires. Elle ne s'applique pas encore aux boursiers ou aux titulaires de chaires. »

### 3.1 Stratégie institutionnelle

TEXTE ORIGINAL : « s'engager à élaborer leurs propres politiques et normes de gestion des données pour les plans de gestion des données »

COMMENTAIRES : Bien que nous reconnaissons la nécessité pour les établissements de répondre aux besoins disciplinaires et administratifs locaux, permettre à chaque établissement d'« *élaborer leurs propres politiques et normes de gestion des données* » n'est pas assez restrictif. Les meilleures pratiques en matière de plans de gestion des données (PGD) sont bien établies – visiter l'[Assistant PGD](#) de Portage pour voir un exemple de la façon dont ces meilleures pratiques sont exprimées dans un outil en ligne. Ces meilleures pratiques devraient établir la norme pour les PGD et éclairer toute politique institutionnelle éventuelle qui est élaborée. Laisser cela aux établissements pourrait entraîner une dilution des exigences du PGD, tant du point de vue du contenu que de la conformité.

TEXTE SUGGÉRÉ :

« s'engager à adopter de meilleures pratiques déjà établies lors de l'élaboration de politiques et de normes institutionnelles pour les plans de gestion des données ».

TEXTE ORIGINAL : « donner et faciliter l'accès aux services de dépôt reconnus ou à d'autres plateformes pour conserver et structurer les données de recherche, et les rendre accessibles à l'avenir. »

COMMENTAIRES : Cela devrait se lire comme suit : « et les rendre accessibles de manière continue et appropriée à l'avenir » pour reconnaître le fait que toutes les données ne peuvent pas être rendues accessibles ouvertement.

TEXTE SUGGÉRÉ :

« ... et les rendre accessibles de manière continue et appropriée à l'avenir »

### 3.2 Plans de gestion des données

#### ***Critères d'admissibilité***

TEXTE ORIGINAL : Toutes les demandes de subvention soumises aux organismes devraient inclure des méthodologies qui tiennent compte des meilleures pratiques en matière de gestion des données de recherche. Les organismes incitent les demandeurs à dresser des plans de gestion des données comme étape essentielle de la conception de projets de recherche. Pour certaines occasions de financement, les organismes subventionnaires peuvent demander que les plans de gestion des données soient soumis à l'organisme approprié au moment de la demande; dans ces cas, les plans de gestion des données peuvent être examinés dans le processus de prise de décision.

COMMENTAIRES : Les exigences en matière de PGD doivent être claires et précises. L'utilisation actuelle des termes « *devrait inclure* », « *incitent* », « *peuvent demander* » et « *peuvent être examinés* » ne répond pas aux attentes, notamment lorsque les PGD sont décrits dans le paragraphe comme étant une « *étape essentielle de la conception de projets de recherche* ». De plus, le préambule de la politique indique que l'utilisation des PGD « *est essentielle à l'excellence en recherche* ». S'ils sont une étape « *essentielle* » et que leur utilisation est également « *essentielle* » alors les exigences en ce qui concerne la soumission, l'évaluation et la mise à jour des PDG prévues dans la politique devraient être renforcées en conséquence.

### **Décision**

COMMENTAIRES : Les trois organismes devraient clarifier leur vision en ce qui concerne la décision du PGD quant au moment où le PGD devrait être présenté et mis à jour (*les PGD sont des « documents évolutifs »*), aux critères utilisés pour évaluer le PGD à diverses étapes dans le cycle de vie de la recherche, et à qui réalisera ces examens. La politique Horizon 2020 de l'UE stipule : « *Le chercheur principal doit élaborer un plan de gestion des données au cours des six premiers mois du projet et le tenir à jour tout au long du projet.* »

TEXTE SUGGÉRÉ :

« Toutes les demandes de subventions soumises aux organismes doivent inclure des plans de gestion de données (PGD) comme étape essentielle de la conception de projets de recherche. Les PGD seront pris en considération dans toutes les demandes de subvention et, dans le cas de possibilité de subventions précises, seront évalués dans le cadre du processus de décision. Les chercheurs principaux des projets approuvés doivent veiller à ce que les PGD soient tenus à jour tout au long du projet. »

## **3.3 Dépôt de données**

### **Définition des données devant être déposées**

TEXTE ORIGINAL : « Les détenteurs de subvention sont tenus de déposer, dans un dépôt numérique reconnu, les données de recherche numérique, les métadonnées et les codes qui appuient directement les conclusions de la recherche dans les publications de revues, des préimpressions et d'autres résultats de recherche qui découlent de la recherche appuyée par les organismes subventionnaires. »

COMMENTAIRES : Ce texte pourrait être utilisé pour limiter considérablement les données qui sont déposées. Ce qui est décrit ici est ce qu'on appelle les « données de réplcation » - le strict minimum de données nécessaires pour reproduire un tableau, un graphique ou une régression (par exemple) dans un article publié.

Tel qu'il est rédigé, les chercheurs pourraient, par exemple, déposer des données pour les cinq variables utilisées dans une régression, ou un graphique. Toutes les

autres variables, qui fournissent du contexte et de la valeur à l'ensemble global de données, pourraient être exclues dans le libellé actuel.

Portage estime qu'il est utile d'élargir la définition pour y inclure toutes les données finales, la documentation et le code associés à un projet de recherche, et pas seulement le sous-ensemble utilisé dans un article publié. Portage croit que c'était l'intention des trois organismes, mais cela doit être précisé dans la politique.

De plus, l'[ESRC](#) (*Economic and Social Research Council*) au R.-U. exige que les titulaires de subvention déposent des données appropriées pour la « *préservation et la réutilisation future* ». Dans le contexte d'un dépôt dans un dépôt numérique, « *appropriées pour la réutilisation* » indique la meilleure pratique d'utiliser des formats non exclusifs, neutres de plateformes qui appuient la préservation, l'archivage et la migration future des données, si la décision de conservation est prise. *Appropriée pour la réutilisation future* témoigne de l'importance de veiller à ce que les données, les métadonnées et les codes déposés dans le dépôt numérique sont complets et de qualité suffisante pour être compréhensibles de façon indépendante par les futurs chercheurs.

TEXTE SUGGÉRÉ :

« Les détenteurs de subvention sont tenus de déposer toutes les données, métadonnées et tous les codes de recherche numériques associés aux conclusions de la recherche financée par l'organisme (c.-à-d. les publications de revues, les préimpressions et autres résultats de recherche) dans un dépôt numérique reconnu, dans un format qui appuie à la fois la préservation et la réutilisation future. »

**Dépôts numériques : « *préservation* » et « *conservation* »**

TEXTE ORIGINAL : « Ce dépôt numérique assurera l'entreposage, la préservation et la conservation sûrs des données. »

COMMENTAIRES : « *entreposage dans un dépôt numérique* » est principalement destiné à la découverte et à l'accès, et en vertu des pratiques exemplaires de GDR, il inclurait certainement la conservation et le stockage sécuritaire des données, de la documentation et du code. Les conservateurs peuvent décider que *certaines* données stockées dans un dépôt devraient être traitées spécialement pour la préservation à long terme dans les *archives*. Dans ce contexte, la *préservation* indique l'application de pratiques et de processus spéciaux qui assurent le stockage et la maintenance à long terme des données, de la documentation et du code, et qui ne sont pas du ressort des dépôts de données.

On se demande ce qu'on entend par « *conservation* » et comment cette exigence s'appliquera aux dépôts. Par exemple, il y a des dépôts qui offrent aux chercheurs des postes « *autonomes* » – sans rôle de conservateur fourni ou assumé. La *conservation* est une entreprise complexe et impliquée pour tout dépôt; il faut

préciser qui assumera ce rôle et comment il sera financé.

TEXTE SUGGÉRÉ :

« Le dépôt numérique assurera la conservation et l'entreposage sûr des données déposées, et facilitera la découverte ainsi qu'un accès approprié. »

***Accès aux données, échéance du dépôt, embargos et relations***

TEXTE ORIGINAL : « Les organismes incitent les chercheurs à donner accès aux données pour autant que les exigences éthiques, juridiques et commerciales le permettent, et conformément aux normes propres à leurs disciplines. Dans la mesure du possible, ces données, ces métadonnées et ces codes doivent être liés à la publication à l'aide d'un identificateur numérique permanent. »

COMMENTAIRES :

Le projet de politique actuel ne mentionne pas quand les chercheurs doivent déposer leurs données. Le document [Concordat on Open Research Data](#) du R-U. affirme que « *les chercheurs rendront, dans la mesure du possible, leurs données de recherche ouvertes et utilisables dans un délai court et bien défini, ce qui peut varier selon le sujet et le domaine disciplinaire et reflète les ressources dont ils disposent à cette fin. Les publications à l'appui des données devraient être accessibles au plus tard à la date de publication et être sous forme de référence. Lorsqu'il n'est pas possible de rendre les données ouvertes pour des raisons légitimes, il ne devrait pas y avoir de conséquences négatives pour les chercheurs concernés.* » L'[ESRC](#) (Economic and Social Research Council) au R-U. exige que les titulaires de subventions déposent des données « *dans les trois mois suivant la fin de la subvention* ».

TEXTE SUGGÉRÉ :

« Les chercheurs rendront, dans la mesure du possible, leurs données de recherche ouvertes et utilisables dans un délai court et bien défini, en fonction de normes disciplinaires. Dans certains cas, le report de l'accès aux données (c.-à-d. une période d'embargo des données clairement définie) peut être justifié. Dans le cas des données qui ne peuvent être entièrement ouvertes pour des raisons éthiques, juridiques, commerciales ou d'autres raisons valables, les métadonnées et le code doivent être ouverts et des ententes d'accès aux données doivent être mises en place pour définir et faciliter un accès contrôlé approprié aux données connexes. Les données, les métadonnées et le code hébergés dans un dépôt devraient être liés aux résultats de la recherche, comme les publications comportant un identificateur numérique persistant. »

#### **4. Date de mise en œuvre**

La collectivité s'inquiète beaucoup du moment et de la façon dont la politique sera mise en œuvre. La planification au niveau institutionnel, régional et national sera influencée par

le calendrier de mise en œuvre de cette politique. Il faut tout mettre en œuvre pour que cela soit rendu public le plus tôt possible.

#### **5. Conformité avec la politique**

TEXTE ORIGINAL : « En acceptant les fonds des organismes subventionnaires, les établissements et les chercheurs acceptent les conditions décrites dans les politiques, les ententes et les lignes directrices des organismes. En cas de violation présumée des politiques, des ententes ou des lignes directrices des organismes, ceux-ci peuvent prendre les mesures prévues pour régler l'allégation, conformément au [Cadre de référence des trois organismes sur la conduite responsable de la recherche](#). »

COMMENTAIRES : Il faut plus d'information ici sur la façon dont la conformité sera surveillée et appliquée. Qui surveillera la conformité et comment? Quelles sont les conséquences de la non-conformité pour un chercheur ou un établissement? Qui paiera pour tout cela? Cette information devrait être énoncée explicitement.

#### **6. Examen de la politique**

Aucun commentaire.

#### **7. Renseignements supplémentaires**

Aucun commentaire.

## Annexe A. Commentaires et questions généraux de la collectivité

### -- basé sur des assemblées locales et d'autres discussions informelles :

- Donner un mandat plus fort aux établissements et aux organismes qui tentent de bâtir l'infrastructure et les services de GDR; nous permet de vendre plus facilement la GDR dans nos établissements et d'obtenir du financement pour appuyer ces initiatives.
- Il sera plus facile pour les établissements de créer des stratégies institutionnelles de GDR et des politiques éventuelles de GDR si les exigences de la politique des trois organismes sont plus rigoureuses.
- Certaines préoccupations au sujet du fait que les établissements au Canada ont des politiques et des normes différentes en matière de GDR – les incohérences pourraient causer des problèmes lorsque des projets de recherche interétablissements et/ou des chercheurs déménagent d'un établissement à un autre.
- L'infrastructure à l'appui de ces initiatives nécessite un investissement dans l'IRN (*et plus précisément la GDR*).
- Les bailleurs de fonds doivent jouer un rôle actif dans l'administration et la conformité à la présente politique.
- Quel financement et quel soutien les trois organismes offriront-ils pour aider les chercheurs et les établissements à se conformer à la politique?
- Nous avons besoin d'un énoncé plus clair sur la façon dont la conformité sera surveillée et appliquée (p. ex. inclure un DOI aux données déposées à la fin d'une subvention?). Qui fera cela et qui paiera?
- Les trois organismes devraient énoncer explicitement (dans la foire aux questions?) l'importance des métadonnées et du code pour appuyer le partage et la réutilisation des données.
- Insister sur l'importance d'exposer les métadonnées et le code même si des restrictions s'appliquent à l'accès aux données réelles.
- Les PGD doivent passer par les trois organismes et faire partie intégrante du cycle de vie des subventions et de la recherche, y compris la nécessité de mettre à jour les PGD pendant la période de subvention.
- Il faut insister davantage sur le fait que les PGD sont des « documents évolutifs » qui doivent être mis à jour tout au long du cycle de vie de la recherche.
- Il faut plus d'information et de clarté sur les questions relatives aux étudiants, aux chaires de recherche du Canada, etc., qui sont exclues de la politique dans sa forme actuelle.
- Nous avons besoin de plus d'information sur le calendrier et le rythme de mise en œuvre de la politique.
- On a manifesté un certain intérêt pour que la politique de GDR s'harmonise plus étroitement avec la politique du gouvernement sur les données ouvertes. Bien qu'il ne s'agisse pas d'une politique de « données ouvertes », la politique devrait mettre davantage l'accent sur la valeur du partage des données.
- Certaines préoccupations au sujet de ce qu'on entend par « dépôt numérique reconnu » – il faut travailler sur l'état de « confiance » (un « sceau d'approbation »?) pour les dépôts.
- Les trois organismes doivent être plus clairs quant au moment où les données doivent être déposées et aux mécanismes permettant de répondre aux préoccupations des chercheurs (p. ex. accords d'accès aux données, embargos).