



# Sommet sur le cadre des services de données nationaux de 2020 Document de travail

rédigé pour le 3<sup>e</sup> Sommet CSDN  
(5 et 6 février, Ottawa, Ontario)

Données de recherche Canada  
*Groupe de travail sur le CSDN, Comité de l'infrastructure*

*Eugene Barsky, Jane Burpee, David Castle, Mark Leggott, Jeff Moon, Karen Payne, Rebecca Ross, Barton Satchwill,  
Diane Sauvé, Byron Spencer, Felicity Taylor, Ela Yazdani*

Version 0.5

## Contexte

L'écosystème de la gestion des données de recherche (GDR) évolue rapidement dans le monde entier, et le Canada n'y échappe pas. Partout, les pays embrassent le concept d'ouverture sous ses multiples facettes, qu'il s'agisse de gouvernement ouvert, de dialogue ouvert, de données ouvertes ou de science ouverte. Les trois conseils subventionnaires élaborent la version définitive d'une politique sur la gestion des données qui donnera plus facilement accès aux résultats de la recherche. Cette nouvelle politique devrait être rendue publique au début de 2020.

La présidence de l'[Open Government Partnership \(OGP\)](#) par le Canada a fait en sorte que notre pays accueille le partenariat à l'occasion de son sommet de 2019, tenu à Ottawa. Le Canada continue aussi d'être un chef de file au sein du groupe des ministres des sciences du G7. Dans leur communiqué de septembre 2017<sup>1</sup>, ces derniers insistaient sur les infrastructures mondiales de la recherche ainsi que sur la science ouverte, et ils appuyaient le travail effectué par le [G7 Open Science Working Group](#), qui met en relief les deux aspects cruciaux que voici :

1. les mesures d'encouragement et l'écosystème de la recherche;
2. les infrastructures autorisant une exploitation optimale des données scientifiques.

Ces priorités et d'autres se reflètent aussi dans les activités visant à définir une nouvelle organisation de l'infrastructure de recherche numérique (NOIRN) qui chapeautera le calcul informatique de pointe (CIP), la gestion des données de recherche (GDR) et les logiciels de recherche (LR). La NOIRN devrait devenir totalement opérationnelle en 2022. La prestation de services IRN robustes et durables aux chercheurs entre donc dans une nouvelle phase palpitante.

Ce document veut servir de toile de fond au Sommet CSDN de 2020 (5 et 6 février) et présente l'approche qui devrait la réalisation des résultats souhaités. Parallèlement, il sert de trait d'union entre deux autres documents récents qui exposent les priorités de l'écosystème de la GDR pour les années à venir.

1. [Data Management Roadmap – 2019-2024](#)<sup>2</sup> (feuille de route de la gestion des données) est un document rédigé par un groupe important d'intervenants à partir des recommandations formulées par le [CUCCIO](#) et présentées à ISDE en 2017<sup>3</sup>. Ce document a eu une influence prononcée sur l'approche actuellement poursuivie pour créer la NOIRN. La feuille de route indique les aspects sur lesquels on devrait se concentrer dans trois domaines : la coordination; l'infrastructure, les outils et les plateformes; les compétences, la constitution d'un réseau d'experts et la communauté

---

<sup>1</sup> [http://www.g7italy.it/sites/default/files/documents/G7%20Science%20Communiqu%c3%a9\\_1/index.pdf](http://www.g7italy.it/sites/default/files/documents/G7%20Science%20Communiqu%c3%a9_1/index.pdf)

<sup>2</sup> Data Management Roadmap – 2019-2024: For Submission to Innovation, Science and Economic Development (ISED). 29 mars 2019 (en anglais seulement)

<sup>3</sup> Élaboré par le groupe de travail sur la gestion des données du Conseil du leadership sur l'infrastructure de recherche numérique. <https://digitalleadership.ca>

de pratique. Ces recommandations portent essentiellement sur la période de transition d'un ou deux ans qui précèdera l'inauguration de la NOIRN. La feuille de route met en relief les secteurs d'activité de plusieurs organisations dont CANARIE, le réseau Portage, DRC et quelques autres. Bien qu'elle marque le consensus d'un petit nombre d'acteurs dans la GD, la collectivité n'a pas participé à sa création.

2. *La gestion des données au Canada : état des lieux* (document d'information sur la GD) [EN / FR] s'inspire également du document que le [CUCCIO](#) avait remis à ISDE en 2017 et fait le point sur l'écosystème de la GDR au Canada. On devrait l'actualiser annuellement ou tous les deux ans. Le document brosse un tableau détaillé de la GDR, notamment des composantes du cycle de la recherche, et décrit quelques acteurs majeurs dans ce domaine. En est toutefois absente une description de l'infrastructure propre à chaque domaine et des communautés de pratique. Le Sommet CSDN de 2020 devrait corriger cela en partie en dressant la liste des acteurs et des services spécifiques aux différents domaines.

La *Déclaration de Kanata (première version)* [EN / FR] est le principal aboutissement du Sommet sur le cadre des services de données nationaux (CSDN) de 2019. Elle reflète le consensus d'un très grand nombre d'intervenants au sujet de ce dont les chercheurs auront besoin au cours des années à venir. Bien que sciemment de nature très générale, la Déclaration exprime le point de vue du plus vaste groupe d'intervenants à ce jour. Elle sera aussi le principal sujet de conversation au Sommet CSDN de 2020 grâce à une version que l'on peut commenter: [EN](#) / [FR](#).

## Peaufiner le CSDN

### Résultats souhaités

Le Sommet CSDN de 2020 amènera la profession à réfléchir sur la feuille de route de la GD, le document d'information et, plus spécifiquement, la Déclaration de Kanata en vue de parvenir aux résultats que voici :

1. Consensus sur les exigences minimales d'un service de données national (SDN)
2. Consensus sur ce qui fonctionne le mieux aux paliers institutionnel, provincial, régional, national et international
3. Consensus sur les services de données nationaux (SDN) existants, automatisés ou pas, que l'on pourrait utiliser comme modèle ou exemple pratique (par ex., CANFAR<sup>4</sup>, Services de gestion des identités de CANARIE<sup>5</sup>, Assistant PGDt<sup>6</sup>);
4. Consensus sur les principales lacunes qui ne figurent pas sur la feuille de route de la gestion des données ou dans la Déclaration de Kanata
5. Consensus sur les priorités pour l'avenir

---

<sup>4</sup> <https://www.canfar.net/fr/>

<sup>5</sup> <https://www.canarie.ca/fr/identite/>

<sup>6</sup> <https://assistant.portagenetwork.ca/fr>

Les résultats du Sommet CSDN de 2020 seront largement diffusés dans l'écosystème de l'IRN canadienne et transmis directement à ISDE ainsi qu'à la jeune NOIRN.

## Modèles

Le Sommet continuera d'utiliser le modèle du nuage européen pour la science ouverte (EOSC)<sup>7</sup> en six volets, mentionné dans la Déclaration de Kanata et repris ci-dessous. Ce modèle a prouvé son utilité à la fois dans le contexte du CSDN canadien et par un meilleur recouplement avec les efforts déployés dans d'autres initiatives nationales.

Architecture	Architecture requise pour créer un système fédéré de plateformes, d'outils et de ressources GD interopérables auxquels tous les chercheurs du Canada auront accès.
Données	Normes et outils communs qui facilitent l'application des principes FAIR dans toutes les disciplines sans interdire l'accès à de riches (méta)données dans un domaine précis.
Services	Services génériques et spécifiques à une discipline soutenant directement les chercheurs et les établissements, de même que les politiques des organismes de financement, des éditeurs et des institutions.
Accès et interfaces	Interfaces humaines et machine qui facilitent le dépôt et la consultation des données tout en garantissant une sécurité robuste et la protection des renseignements personnels, s'il y a lieu.
Règles	Jeux de politiques et de processus qui favorisent la participation de tous les membres de l'écosystème de la recherche, soutiennent les différentes juridictions et suscitent la confiance.
Gouvernance	Cadre faisant en sorte que tous ceux qui jouent un rôle dans le développement et la durabilité de l'infrastructure de recherche numérique et de l'écosystème de la recherche canadiens soient correctement représentés.

Figure 1P : Le nuage européen pour la science ouverte (adapté pour le Canada)

La collectivité canadienne recourt aussi à un modèle d'architecture générale pour orienter le dialogue et les activités sur le CSDN, qu'il s'agisse d'alimenter le débat ordinaire ou de répondre à des appels à projets spécifiques. Le Sommet CSDN de 2020 permettra à la collectivité d'étoffer le modèle de CSDN illustré plus bas.

<sup>7</sup> <http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

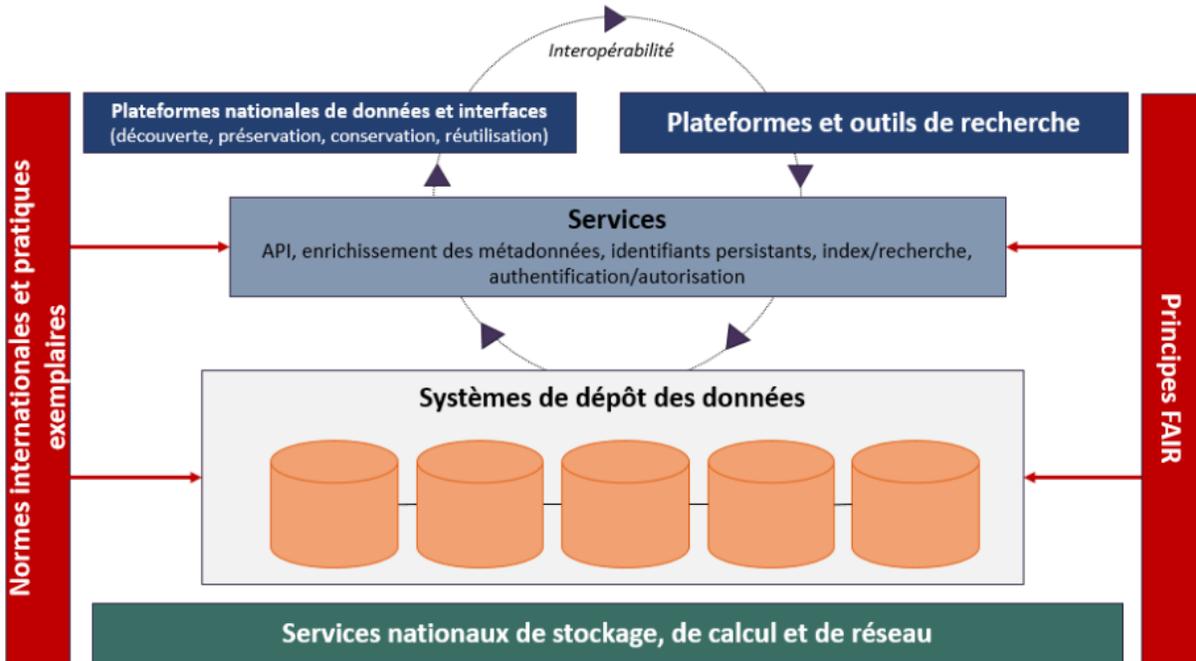
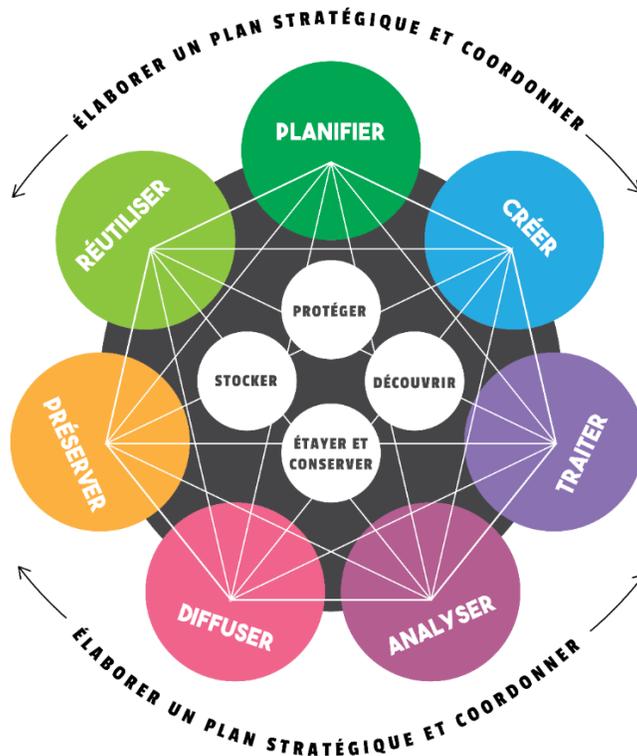


Figure 2. Modèle du CSDN de Données de recherche Canada

Le cycle de la recherche illustre les étapes par lesquelles passe un projet de recherche et rend la tâche plus facile aux personnes et aux institutions qui entretiennent des liens directs avec les chercheurs. Le diagramme du cycle de la recherche présenté ci-dessous est tiré du document présenté par le [CUCCIO](#) à ISDE en 2017. À l’instar de nombreux modèles similaires, le cercle extérieur illustre les étapes principales du processus de recherche et ses points d’intersection avec les services qui ont un impact sur lui, à chacune d’elles. Le diagramme qui suit utilise quatre services, mais il peut y en avoir davantage, certains propres à un domaine.



## Prochaines étapes

Le Sommet CSDN de 2020 réunit plusieurs conférenciers et prévoit trois exercices de réflexion en petits comités, de même qu'un mécanisme pour valider et regrouper les résultats des délibérations. Dans les exercices de réflexion, six à huit personnes se concerteront pour répondre aux interventions précises ayant pour but de mettre à jour la version initiale de la Déclaration de Kanata et de parvenir aux résultats énumérés plus haut. Il en découlera une version plus actuelle de la Déclaration et une liste mieux détaillée des SDN.

Avant le Sommet, les membres de la profession GDR auront l'occasion d'annoter la première version de la Déclaration de Kanata en formulant leurs commentaires, en soulignant ses lacunes et en énumérant les SDN existants. Tout cela sera intégré à l'ébauche de la nouvelle version qu'on diffusera deux ou trois semaines avant le Sommet et qui servira de point de départ aux rencontres en petits groupes lors du Sommet. Deux ou trois mois après ce dernier, les résultats seront largement diffusés à la collectivité GD.