



# Recommandations pour assurer le bon fonctionnement d'un centre d'accès aux données pour la recherche au sein d'un organisme public

Version 2 – Juin 2022

## Introduction

Au cours des dernières années, l'importance des données collectées par les organismes publics du Québec s'est affirmée avec force dans les discours et les pratiques. Qu'elles soient de nature sociale, économique, fiscale ou clinique, les données constituent des actifs collectifs qui sont indispensables à la poursuite de la recherche et à l'avancée des connaissances scientifiques. Il est donc essentiel de favoriser leur valorisation au service de l'intérêt public tout en garantissant leur protection et leur gestion responsable.

L'objectif de ce document est de fournir des recommandations en vue de guider les organismes publics du Québec dans le traitement des demandes d'accès aux données émanant des communautés de recherche<sup>1</sup>. Avec les nouvelles avancées législatives et réglementaires, les organismes publics acquièrent désormais de nouveaux rôles et responsabilités en matière d'accès aux données. Plusieurs d'entre eux sont et seront amenés à mettre en place des centres d'accès aux données (aussi appelés "guichets d'accès" ou "portes d'accès") pour traiter les demandes des chercheuses et des chercheurs. Il est donc indispensable que les organismes publics adoptent les meilleurs standards et pratiques, de manière à répondre aux demandes d'accès aux données avec efficacité et responsabilité. Ceci contribue activement à la construction d'un lien de confiance durable avec la communauté de recherche et le public.

Dans cet objectif, le présent document se propose d'outiller les organismes publics dans le bon fonctionnement de leurs centres d'accès aux données en vertu de valeurs et de principes centraux, tels que la transparence, l'équité, la conduite responsable en recherche, la science ouverte et la reproductibilité des connaissances scientifiques. Il s'agit là de valeurs et de principes incarnés et promus par les Fonds de recherche du Québec (FRQ) à travers leurs missions et leurs actions au service de la communauté de recherche, ainsi que des citoyennes et des citoyens.

En octobre 2021, les FRQ ont formé un *Comité pour des données accessibles à la recherche*. Sa mission principale est d'analyser la situation actuelle de l'accès aux données à des fins de recherche et de proposer des améliorations concrètes pour favoriser la production et la

---

<sup>1</sup> Par organismes publics, nous entendons les organismes tels que les ministères, instituts, centres et établissements qui sont financés pour une bonne part par l'État et dont la majorité des dirigeants sont nommés par le gouvernement ou par l'un de ses ministres.



diffusion des connaissances scientifiques. Le présent document est le fruit des échanges et réflexions des membres du Comité. Il ne s'agit pas d'un produit final mais bien d'un travail en continu qui demeure constamment nourri et adapté en fonction des modifications législatives et réglementaires, ainsi que des nouveaux standards de pratiques, politiques et procédures de référence en matière d'accès et de gestion des données à des fins de recherche. Finalement, ce document n'a pas de portée prescriptive mais se présente plutôt comme un guide de bonnes pratiques pour soutenir les organismes publics dans la mise en œuvre de leurs nouvelles missions au service de la recherche et de la progression de la science ouverte.



## *Le Comité des FRQ pour des données accessibles à la recherche*

Le Comité des FRQ pour des données accessibles à la recherche est composé de 26 membres issus des communautés de recherche du Québec et du Canada, ainsi que des ministères et organismes publics du Québec. Ce groupe de travail s'articule autour d'un comité exécutif, porteur de propositions et recommandations initiales, et d'un comité élargi, apportant l'expertise nécessaire à leur validation et enrichissement.

### **Comité exécutif**

Anick Bérard (Université de Montréal et CHU Ste-Justine)  
Julie Couture (Ministère de la Santé et des Services Sociaux)  
Mariève Doucet (Ministère de la Santé et des Services Sociaux)  
Pierre Doray (Université du Québec à Montréal)  
Jason Robert Guertin (Université Laval)  
Anaïs Lacasse (Université du Québec en Abitibi-Témiscamingue)  
Alexandra Obadia (CARTaGENE)  
Jean-Pierre Perreault (Université de Sherbrooke)  
Robert Platt (Université McGill)  
Mylène Deschênes (Fonds de recherche du Québec)  
Carole Jabet (Fonds de recherche du Québec - Santé)  
Cécile Petitgand (Fonds de recherche du Québec - Santé et Centre de recherche du Centre hospitalier de l'Université de Montréal)

### **Comité élargi**

James Brophy (Université McGill)  
Sylvana Côté (Université de Montréal)  
Philippe Després (Université Laval)  
Sherif Eltonsy (Université du Manitoba)  
Deshayne Fell (Université d'Ottawa)  
Antoine Groulx (Université Laval)  
Catherine Haeck (Université du Québec à Montréal)  
Padma Kaul (Université de l'Alberta)  
Jacinthe Leclerc (Université du Québec à Trois-Rivières)  
Michael Paterson (ICES)  
Sylvie Perreault (Université de Montréal)  
Nadia Sourial (Université de Montréal)  
Robyn Tamblyn (Université McGill)  
Alain Vanasse (Université de Sherbrooke)

⇒ [Lien vers le mandat du Comité](#)



## Table des matières

Introduction .....	1
Le Comité des FRQ pour des données accessibles à la recherche .....	3
Table des matières .....	4
1. Préparation de la demande d'accès aux données.....	5
1.1. Catalogue des banques de données et variables.....	5
1.2. Processus d'accès aux données.....	6
1.3. Évaluation des coûts.....	7
1.4. Trousse à outils pour les chercheuses et les chercheurs .....	8
1.5. Exemples de demandes types .....	9
2. Dépôt et suivi de la demande d'accès aux données.....	9
2.1. Espace projet pour le dépôt et le suivi.....	9
2.2. Canal de communication dédié .....	10
2.3. Traçabilité et reproductibilité .....	10
3. Accès aux données après autorisation.....	11
3.1. Modalités d'accès aux données .....	11
3.2. Outils disponibles et services offerts .....	12
3.3. Documentation complémentaire.....	13
4. Partage des résultats et valorisation de la recherche .....	13
4.1. Principes FAIR et science ouverte .....	13
4.2. Diffusion des résultats auprès du public.....	14
4.3. Transparence, partage et collaboration.....	15
5. Responsabilités du Centre d'accès .....	16
5.1. Informations sur le fonctionnement et les personnes-ressources .....	16
5.2. Protection de la sécurité des données .....	16
5.3. Indicateurs de performance .....	17
5.4. Registre des projets de recherche .....	18



## 1. Préparation de la demande d'accès aux données

### 1.1. Catalogue des banques de données et variables

Sur son site Internet, le Centre d'accès devrait mettre à disposition des chercheuses et des chercheurs un catalogue exhaustif listant :

- **Toutes les banques de données disponibles via le Centre**

Ceci inclut les jumelages et appariements de données déjà réalisés (à partir des banques de données du Centre, et entre les données du Centre et celles des chercheuses et chercheurs). Il est aussi important d'indiquer quels jumelages et appariements de données sont autorisés (à partir des banques de données du Centre, et entre les données du Centre et celles des chercheuses et des chercheurs).

- **Les variables contenues dans chaque banque de données et leur domaine de valeur associé**

Il est essentiel d'explicitier la codification de la variable et la signification du code, et aussi, lorsque possible, de fournir des liens vers des sources externes permettant de valider l'information fournie.

Idéalement, les variables liées sont listées en utilisant des concepts reposant sur des terminologies/ontologies reconnues dans chaque domaine (par exemple ICD-11, LOINC ou SNOMED CT dans le domaine clinique).

Les informations concernant les banques de données et les variables (appelées « métadonnées ») devraient être accessibles dans un format ouvert courant et facilement lisible par la chercheuse ou le chercheur, et même un algorithme, tel que le format .csv ou xml.

**Intéressant** : La [plateforme Maelstrom](#) de l'Université McGill a contribué au développement de la suite logicielle libre OBiBa pour les centres d'accès aux données et la communauté de recherche en vue de faciliter l'adoption de standards (tels que HL7 FHIR) pour favoriser le jumelage et l'échange de données. À cet égard, la Plateforme PULSAR de l'Université Laval a adopté la suite OBiBa pour gérer les données, les documenter et générer des catalogues de métadonnées, tel qu'illustré par l'[étude NutriQuébec](#).

Le Centre d'accès pourrait aussi rendre disponibles des **fichiers de données synthétiques** (jeux des données fictives) afin de permettre aux chercheuses et aux chercheurs de conduire des analyses exploratoires et de mieux planifier leur demande d'accès aux données. À ces fichiers de données synthétiques devraient être associées des métadonnées précisant les méthodes employées pour générer ces données et les variations statistiques existantes vis-



à-vis des données réelles. Les fichiers de données synthétiques, enrichis par ces métadonnées, permettent ainsi aux chercheuses et aux chercheurs d'écrire leur code et de conduire leurs premières analyses à distance afin de venir au Centre d'accès une seule fois pour obtenir leurs résultats à partir des données réelles.

💡 **Intéressant** : Les chercheuses et les chercheurs qui demandent l'accès à la [cohorte TorSaDE](#) à travers l'Institut de la statistique du Québec peuvent recevoir les variables de la cohorte avec des données fictives et aléatoires. Ceci s'avère très utile pour effectuer des analyses tests et adapter leurs questions de recherche et les objectifs visés avant la soumission de la demande d'accès. Idéalement, la cohorte fictive pourrait même être partagée à d'autres chercheuses et chercheurs pour servir la formation d'étudiantes et d'étudiants gradués et d'autres membres de la communauté de recherche travaillant sur ce même type de données.

Aux États-Unis, les [Centers for Medicare & Medicaid Services](#) fournissent aux chercheuses et aux chercheurs des jeux de données synthétiques afin de conduire des analyses en amont de leur demande d'accès aux données. Au Royaume-Uni, le [Clinical Practice Research Datalink \(CPRD\)](#) propose à la communauté de recherche plusieurs types de jeux de données synthétiques pour l'analyse statistique et le calibrage d'algorithmes d'apprentissage automatique.

### 1.2. Processus d'accès aux données

Le Centre d'accès devrait décrire sur son site Internet les principaux processus d'accès aux données que doivent parcourir les chercheuses et les chercheurs. Ces processus peuvent varier en fonction de : l'affiliation du demandeur d'accès aux données, le type de banques de données ou de variables requis, la nature du projet de recherche, le type d'autorisation exigée, le consentement obtenu pour l'utilisation des données (par exemple pour les données de cohorte), entre autres facteurs. Il s'agit en fait pour le Centre d'accès de baliser les principaux cheminements que suivent les demandes d'accès aux données.

Chaque type de processus d'accès décrit sur le site Internet du Centre devrait contenir minimalement les informations suivantes :

- Les caractéristiques principales du processus : type de chercheuse ou de chercheur concerné, type de données requises, type de projet, type d'autorisation exigée, si les données ont été consenties pour utilisation secondaire ou non, etc. ;
- Les principales étapes du processus d'accès, incluant les signatures d'ententes et de contrats ;
- Les délais moyens requis par chaque étape et les délais minimal et maximal pour l'ensemble du processus ;
- Un lien vers la documentation exigée pour compléter les étapes du processus (voir **1.4.** ci-dessous).



La description d'un processus d'accès standard peut être complétée par une infographie permettant aux chercheuses et aux chercheurs de visualiser clairement et simplement les principales caractéristiques et étapes du processus en un coup d'œil.

💡 **Intéressant** : En Ontario, ICES a élaboré [un schéma assorti d'explications précises et concises](#) pour présenter les étapes du processus standard d'accès aux données. Les explications apportées indiquent notamment les actions à accomplir par les employés d'ICES d'une part, et par les chercheuses et les chercheurs d'autre part.

Le Réseau de recherche sur les données de santé du Canada, visant à faciliter la conduite d'études multi-juridictionnelles à travers le pays, définit de façon très claire les principales étapes du processus d'accès aux données [sur son site Internet](#). On y indique notamment par des hyperliens les documents exigés des chercheuses et des chercheurs pour préparer la demande d'accès.

Vous pouvez aussi visualiser les [infographies élaborées par le Health Data Hub](#), plateforme française pour l'accès aux données de santé, pour illustrer les principaux processus d'accès selon les types de demandes et de projets.

**Point important** : Dans le cadre du processus de traitement des demandes, le Centre d'accès ne devrait pas se prononcer sur la pertinence scientifique d'un protocole de recherche qui a fait l'objet d'une évaluation préalable d'un organisme subventionnaire et/ou a reçu une autorisation de la part d'un Comité d'éthique à la recherche. Le Centre d'accès devrait reconnaître que ce type d'évaluation scientifique n'est pas de son ressort et prendre en compte les autorisations données par les organismes de référence en matière d'évaluation et de validation de la recherche.

### 1.3. Évaluation des coûts

Le Centre d'accès devrait exposer de façon transparente les coûts associés aux activités d'extraction, de manipulation et d'analyse des données qui sont conduites en son sein. Si les coûts varient en fonction du statut des demandeurs ou du type de demande, ces informations doivent être précisées sur le site Internet du Centre. Notamment, il est nécessaire de préciser les échelles de coûts pour les chercheuses et les chercheurs des organismes publics/académiques, en comparaison avec ceux pour la recherche financée par un organisme à but non lucratif ou une compagnie privée.

De plus, les coûts associés à l'extraction et la préparation des fichiers de données devraient être moindres si l'équipe de recherche souhaite réutiliser des banques de données existantes (constituées pour d'autres projets) ou les mettre à jour en gardant les mêmes bases de données ou variables. Pour les équipes de recherche sollicitant la même stratégie d'extraction



de données (pour ajouter des informations à une cohorte par exemple) après avoir eu accès aux données précédemment, les coûts du Centre d'accès devraient être d'autant plus réduits.

Aussi, sur le site du Centre d'accès, un simulateur de coûts devrait permettre aux chercheuses et aux chercheurs de pouvoir élaborer un budget en amont de leur demande, en fonction du type de données ou de variables requises pour leur projet. La chercheuse ou le chercheur peut ainsi générer en temps réel plusieurs simulations afin d'obtenir des estimations de coûts pour différentes versions de son projet.

De surcroît, toute soumission envoyée par le Centre, avant ou après l'approbation de la demande d'accès, devrait indiquer clairement à quelle(s) activité(s) le coût estimé correspond (extraction, manipulation, analyse, soutien, etc.) ainsi que l'estimation du nombre d'heures de travail qui seront réalisées par les experts et expertes du Centre d'accès en réponse à la demande de l'équipe de recherche. Ceci contribue dès lors à la transparence des pratiques vis-à-vis de la communauté.

**Intéressant** : L'Institut de la statistique du Québec met à disposition des chercheuses et des chercheurs un simulateur de coûts afin de les aider à explorer la faisabilité de leur projet avant la soumission d'une demande complète. Ce simulateur se trouve dans la Zone chercheur dédiée au suivi de la demande d'accès.

#### 1.4. Trousse à outils pour les chercheuses et les chercheurs

Le Centre d'accès devrait fournir une trousse à outils pour guider les chercheuses et les chercheurs dans la préparation de leur demande d'accès aux données.

**Intéressant** : Le Health Data Hub de France propose un [guide de démarrage](#) pour les chercheuses et les chercheurs qui débutent une procédure de demande d'accès aux données.

La trousse à outils mise à disposition par le Centre d'accès regroupe l'ensemble des informations et des documents à fournir en vue de l'évaluation de la demande d'accès aux données. Cette trousse contient minimalement :

- Le document permettant d'identifier la chercheuse ou le chercheur, ainsi que son organisation d'affiliation principale (en cas de changement récent de la direction générale de l'organisation, il est utile de l'indiquer au Centre et de justifier ce changement) ;
- Le formulaire permettant de collecter des informations sur la nature de la demande d'accès aux données : banques et variables demandées ;
- Une présentation détaillée des activités liées au projet de recherche ;
- Une évaluation des facteurs relatifs à la vie privée (EFVP) ;



- L'approbation du projet par un comité d'éthique de la recherche (lorsqu'applicable, étant donné qu'elle n'est pas indispensable pour tous les projets, par exemple pour un usage secondaire des données de Statistiques Canada) ;
- L'entente de communication entre la chercheuse ou le chercheur et le Centre d'accès à compléter et signer (entente découlant d'une [obligation de la Loi 25](#) pour les données des organismes publics qui ne sont pas des renseignements de santé et de services sociaux) ;
- Les modèles de contrats à signer entre la chercheuse ou le chercheur, le Centre d'accès et toute autre personne ou entité pertinente dans le cadre la demande d'accès ;
- Un guide de démarrage pour orienter les chercheuses et les chercheurs dans une première demande ;
- La liste des personnes-ressources (nom et contact) du Centre en cas de besoin ;
- Le lien vers la Foire aux questions du Centre.

**Intéressant** : La liste de documents qu'un Centre peut exiger pour autoriser la demande d'accès varie en fonction du type de données nécessaires à la réalisation du projet de recherche. Les principaux scénarios d'accès aux données, ainsi que les documents exigés en fonction des lois applicables, sont analysés [sur ce lien](#).

### 1.5. Exemples de demandes types

Afin d'aider les chercheuses et les chercheurs dans la préparation de leur soumission, le Centre d'accès pourrait leur fournir des exemples de demandes (fictives ou inspirées de cas réels anonymisés) qui remplissent les critères exigés et correspondent à des cas de succès dans le processus d'accès aux données. Dans ces exemples, le Centre pourrait rendre explicites les causes et manquements pouvant conduire à un refus de la demande d'accès (par exemple : une documentation incomplète ou l'absence d'information concernant le financement de la recherche). L'équipe de recherche pourra ainsi mieux comprendre les attentes du Centre en matière de remplissage des documents et des modèles de la trousse à outils en vue de l'évaluation de la demande d'accès et de l'efficacité de son traitement.

## 2. Dépôt et suivi de la demande d'accès aux données

### 2.1. Espace projet pour le dépôt et le suivi

Le Centre d'accès devrait permettre aux chercheuses et aux chercheurs de disposer d'un espace en ligne dédié à leur projet dans lequel toutes les informations y afférant sont consignées, et ce du moment du dépôt de la demande à la clôture du projet.



Ainsi, c'est dans l'espace projet que la chercheuse ou le chercheur dépose l'ensemble des documents requis par sa demande, avant et après approbation. Et c'est aussi là qu'elle ou il peut suivre l'avancement de sa demande. À ce titre, un système de signalement (de type voyant vert-orange-rouge ou validé-en cours-refusé) lui permet de savoir à quelle étape du processus d'accès elle ou il se trouve, quels sont les points d'achoppement rencontrés, ou encore quels documents ou informations manquent à son dossier afin d'être complet. À cet égard, un délai maximal de traitement devrait être indiqué pour chaque étape du processus d'accès, afin d'informer si la demande de la chercheuse ou du chercheur accuse un retard et pour quelles raisons.

De plus, lorsque le Centre d'accès est doté d'un système d'information dédié à la gestion des projets de recherche et des processus d'approbation éthique (tel que Nagano par exemple), l'espace projet pour le dépôt et le suivi de la demande d'accès aux données devrait idéalement être connecté à ce système d'information. Une telle connexion entre systèmes permet de gagner en efficacité et en agilité pour assurer la gestion et le suivi de toutes les informations relatives à un projet de recherche fondé sur l'utilisation de données.

## *2.2. Canal de communication dédié*

Au sein de l'espace projet accessible par la chercheuse ou le chercheur, un canal de communication dédié devrait être installé afin de lui permettre de communiquer avec l'équipe du Centre chargée de l'évaluation et du suivi de la demande d'accès.

Dès lors, les communications par courriels entre l'équipe du Centre et l'équipe de recherche devraient être réduites au plus strict minimum ou aux cas urgents. Toutefois, dès qu'une nouvelle information est entrée dans l'espace projet, la chercheuse ou le chercheur devrait recevoir une notification automatique par courriel afin de pouvoir faire un suivi rapide.

Au sein de l'espace projet, toutes les communications devraient être enregistrées pour être accessibles en tout temps par la chercheuse ou le chercheur, ainsi que par l'équipe du Centre d'accès, depuis le moment du dépôt de la demande jusqu'à la clôture du projet.

## *2.3. Traçabilité et reproductibilité*

Toute action effectuée par l'équipe du Centre d'accès pour extraire ou préparer les fichiers de données en réponse à la demande devrait être retraçable, analysable et explicable. L'intégralité du code qui a servi à extraire et préparer les données pourrait même être consignée dans un dépôt approprié (par exemple, GitLab) pour références futures. Ce dépôt deviendrait ainsi une bibliothèque d'outils communs pour la recherche et un gage de transparence supplémentaire de la part du Centre. Une chercheuse ou un chercheur pourrait alors savoir précisément à quel type de travail sur les données est associé chacun des coûts engagés par le Centre dans le cadre du traitement d'une demande d'accès.



L'objectif de ce suivi est finalement de garantir la possibilité pour la chercheuse ou le chercheur demandeur de données, et plus largement toute la communauté de recherche intéressée par le projet, de réviser les données mises à disposition par le Centre. L'équipe de recherche devrait par ailleurs connaître la personne contact du Centre dédiée au projet ou à la demande d'accès (la personne qui a fait l'extraction de données et non un gestionnaire). Il s'agit là de contribuer à la reproductibilité des résultats de la recherche en permettant à la communauté de retracer le cycle de vie complet de la donnée, de sa collecte à sa diffusion, en passant par sa préparation et sa valorisation, en ayant bien connaissance de la durée et des caractéristiques de chacune des étapes du processus.

### 3. Accès aux données après autorisation

#### 3.1. Modalités d'accès aux données

Le Centre d'accès devrait fournir à la chercheuse ou au chercheur toutes les informations relatives aux modalités d'accès aux données (sur place ou à distance, pourquoi et comment) en amont de sa demande. Idéalement, ces modalités d'accès ne devraient pas changer au cours du processus d'accès ni à la suite d'une décision prise par le Centre d'accès sans consultation de l'équipe de recherche. Si l'accès doit s'effectuer par une présence physique dans le Centre, au sein d'un espace sécurisé, la chercheuse ou le chercheur devrait en être informé en amont de l'autorisation et connaître l'ensemble des procédures et protocoles associés à ce mode d'accès, et les raisons justifiant un tel niveau de sécurité. Les contraintes d'accès reliées soit au statut de la chercheuse ou du chercheur, ou de son personnel de recherche, soit au type de son institution d'affiliation devraient être fournies. Ces informations devraient en outre être consignées au sein d'un document accessible par les chercheuses et les chercheurs, ainsi que par le public. Ce document permet aussi de mettre en évidence les procédures et processus en place pour garantir la protection de la sécurité et de la confidentialité des données.

**Intéressant** : L'Institut de statistiques du Québec a produit des [documents explicatifs à l'intention des chercheuses et des chercheurs](#) afin de décrire les modalités d'accès aux données via le Centre d'accès aux données de recherche de l'Institut de statistiques du Québec (CADRISQ).

Dans le cas où l'accès aux données est permis à distance, via un canal sécurisé, le Centre devrait informer la chercheuse ou le chercheur de l'ensemble des exigences associées à cette modalité d'accès, et ce en amont du dépôt de la demande. À ce titre, les exigences en matière de sécurité des ordinateurs utilisés pour accéder aux données à distance devraient être clarifiées d'emblée pour les équipes de recherche, qui pourront ainsi fournir au Centre d'accès les spécificités de leurs dispositifs au moment du dépôt de la demande.



De surcroît, la chercheuse ou le chercheur devrait savoir en amont de ses démarches si l'accès à distance a un impact sur le type de fichiers de données ou de variables accessibles par ce biais. Si les fichiers de données sont masqués ou modifiés d'une quelconque manière du fait d'un accès à distance, cette information devrait être transmise à la chercheuse ou au chercheur avant le processus d'accès.

**Dans tous les cas, l'accès à distance à des fichiers de données non masqués, tel que l'exige la conduite d'un projet de recherche, devrait constituer la solution privilégiée par le Centre lorsque le niveau de sensibilité des données accédées le permet.**

**Intéressant** : StatCan a établi une échelle multi-niveaux standardisée pour déterminer les modalités d'accès aux données (à distance, dans un espace sécurisé, dans les locaux de StatCan) en fonction du statut du demandeur, du type d'organisation et du niveau de sensibilité des données. Ce type d'accès multi-niveaux est décrit dans cette présentation accessible sur [le site Internet du Centre interuniversitaire québécois de statistiques sociales \(CIQSS\)](#).

### 3.2. Outils disponibles et services offerts

Que ce soit sur place ou à distance, via un accès sécurisé, le Centre d'accès devrait fournir aux chercheuses et aux chercheurs une panoplie d'outils d'analyse des données adaptés aux différents domaines de la recherche et évolutifs en fonction de ses nouvelles méthodes et pratiques. Ces outils incluent des logiciels tels que SAS, Stata, R, Python, dont les dernières versions stables doivent être rendues disponibles aux chercheuses et aux chercheurs. Des codes et des scripts, consignés dans un dépôt approprié (voir **2.3.**) peuvent aussi être mis à disposition de la communauté de recherche pour faciliter la valorisation des données. De plus, le Centre d'accès doit pouvoir faciliter la création, l'évaluation et la modification de scripts nouveaux, comme le requiert la recherche en biostatistique et en apprentissage automatique. Finalement, la liste des outils disponibles à travers le Centre devrait être rendue publique sur le site Internet du Centre afin d'en informer la communauté de recherche en amont de toute demande d'accès aux données.

Le Centre d'accès peut aussi proposer des services de stockage et d'analyse de données en permettant aux équipes de recherche d'entreposer en son sein leurs propres fichiers de données en vue d'une valorisation par l'équipe ou par les experts du Centre. De la sorte, le Centre peut choisir d'offrir des capacités de calcul informatique de pointe (pour des projets types en intelligence artificielle, par exemple) ainsi que des services d'analyse au besoin, moyennant le versement d'une contribution supplémentaire par l'équipe de recherche. Le montant de cette contribution doit être connu par avance par les chercheuses et les chercheurs.



### 3.3. Documentation complémentaire

Au cours du processus d'accès aux données, des documents ou des informations supplémentaires pourraient être demandés par le Centre à la suite d'une demande de modification initiée par la chercheuse ou le chercheur. Il peut s'agir par exemple de l'ajout d'un nouveau membre à l'équipe de recherche ou bien d'une entente nouvelle entre l'équipe et une autre institution dans le cadre du projet de recherche. Dans tous les cas, la documentation complémentaire devrait toujours être communiquée via le canal de communication dédié de l'espace projet (voir 2.2. à ce titre). **Et elle ne peut émaner unilatéralement du Centre lui-même ; il s'agit d'une réponse à une nouvelle demande de l'équipe de recherche.**

**Point important :** Pour l'ensemble des demandes d'accès, la documentation indiquée dans la trousse à outils des chercheuses et des chercheurs, pour un processus d'accès en particulier, devrait demeurer prévisible et inchangée à travers le processus, de telle sorte que la demande de documents complémentaires devrait relever de l'exception plutôt que de la règle.

## 4. Partage des résultats et valorisation de la recherche

### 4.1. Principes FAIR et science ouverte

En vertu des principes FAIR (Facilement trouvable, Accessible, Interopérable et Réutilisable), le Centre devrait favoriser, en collaboration avec les équipes de recherche, la diffusion des résultats de la recherche produits à partir des données accessibles via son processus d'accès. Ces résultats de la recherche incluent les fichiers de données (ou du moins les métadonnées associées), les variables ainsi que les codes développés par la chercheuse ou le chercheur à partir des données mises à disposition par le Centre. Les principes FAIR n'impliquent pas d'emblée que tous les résultats de la recherche doivent être rendus publics sans conditions, mais que ceux-ci puissent être découvrables, accessibles, interprétables, partageables et réutilisables par la communauté de recherche. À cet égard, le Centre pourrait guider les chercheuses et les chercheurs dans l'entreposage de leurs résultats de recherche dans des dépôts institutionnels reconnus tels que le [Dépôt fédéré de données de recherche](#) (DFDR). Ce type de pratique au service de la science ouverte favorise la reproductibilité des connaissances scientifiques et sert la croissance et le partage des savoirs.

**Intéressant :** Le [FAIR Evaluation Services](#), un outil en ligne entièrement gratuit, permet d'évaluer si les ressources d'un Centre d'accès aux données respectent les principes FAIR. Suite à une évaluation, un fichier de données est noté selon chacune des composantes des principes FAIR à travers une méthode de voyants vert et rouge, comme le montre [l'évaluation conduite sur le fichier de données NutriQuébec](#) mis à disposition par la plateforme PULSAR de l'Université Laval.



Au niveau international, le consortium [FAIRsharing](#) a développé un ensemble de guides, outils et services pour venir soutenir les centres d'accès aux données et les communautés de recherche dans l'adoption de pratiques exemplaires en matière de description, documentation, structuration, utilisation et partage de données en vertu des principes FAIR.

Le principe de science ouverte implique notamment l'établissement d'un registre des publications disponible pour la communauté et le public. Ainsi, le Centre d'accès devrait mettre à disposition de toutes et tous une liste toujours à jour des publications issues de la valorisation des données mises à la disposition des équipes de recherche. À cet égard, toute chercheuse ou tout chercheur ayant eu accès aux données du Centre devrait recevoir un ensemble de libellés multilingues qu'elle ou il devrait intégrer à ses publications afin de reconnaître la contribution du Centre en tant que source de données. Ce libellé devrait aussi indiquer les modalités d'utilisation des données et des résultats de la recherche en fonction notamment des licences permises par le Centre d'accès.

#### 4.2. Diffusion des résultats auprès du public

Pour informer le public des résultats de la recherche conduite grâce aux données du Centre, la création d'un registre des fichiers, des variables, des codes mais aussi des publications est certes une condition nécessaire à la valorisation des connaissances scientifiques, mais elle n'est pas suffisante. Ainsi, toute publication contenue dans le registre du Centre devrait être dotée d'un résumé vulgarisé rédigé par les chercheuses et les chercheurs et visant à décrire le projet de recherche en des mots simples et accessibles pour les non-initiés aux méthodes scientifiques. Ces textes, décrivant brièvement le processus de la recherche, les données utilisées, les résultats obtenus et leur "impact social", ne devraient pas exiger une formation scientifique préalable pour être compris et devraient être rendus disponibles en français ou en version bilingue.

Outre la constitution d'un registre des publications, il est important pour le Centre d'accès d'accompagner les chercheuses et les chercheurs dans la diffusion de leurs résultats auprès de la communauté, et ce en utilisant une panoplie d'outils éprouvés en communication scientifique. Ces outils permettent en effet de soutenir la compréhension et l'appropriation des méthodes et des résultats de la recherche par les citoyens et les citoyennes. Parmi ces outils, citons les récits d'impact, infographies, balados-diffusion et vidéos décrivant les étapes et les bénéfices d'un projet de recherche, ou encore les plateformes et événements visant à encourager le dialogue entre scientifiques et membres de la société civile.

L'ensemble de ces actions en faveur de la valorisation des résultats de la recherche permettront ainsi de démontrer concrètement l'intérêt public de la recherche fondée sur les données. De la sorte, le Centre d'accès devrait considérer ces modes de diffusion des résultats de la recherche au même titre que les publications scientifiques et les intégrer à son registre des projets de recherche (voir **5.4.** à ce titre).



💡 **Intéressant** : En Ontario, ICES propose des [infographies claires et concises](#) pour mettre de l'avant les résultats de certains projets utilisant les données transitant via cet organisme public au service de la recherche.

En France, le Health Data Hub a produit [une vidéo](#) pour décrire le parcours patient associé à un projet de valorisation des données visant à développer un algorithme d'intelligence artificielle pour mieux traiter les patients porteurs de stimulateurs cardiaques. Le Health Data Hub produit aussi [des posters](#) pour décrire en des termes clairs ses bases de données et rendre compte des utilisations qui en sont faites pour la recherche. Ce type de représentation peut servir d'exemple pour élaborer un résumé de publication qui soit plus accessible au grand public.

#### 4.3. *Transparence, partage et collaboration*

Le Centre d'accès devrait finalement permettre à une chercheuse ou un chercheur ayant achevé son projet de conserver au sein du Centre son jeu de données de recherche ainsi que le code utilisé pour l'analyse de données. Plutôt que de le détruire, le Centre pourrait permettre à la chercheuse ou au chercheur d'y avoir accès ultérieurement dans le cadre d'un autre projet et à travers le dépôt d'une nouvelle demande d'accès. Un tel processus permet ainsi une valorisation optimale des données et une utilisation efficiente des ressources engagées pour les préparer et les valoriser à des fins de recherche. Toute alternative à la destruction des données devrait donc être recherchée par le Centre en vue de garantir l'application des Principes FAIR (embrassés notamment par [les Trois-Conseils au fédéral](#)) et de contribuer à l'optimisation des ressources publiques engagées dans l'extraction, la préparation et l'analyse des données.

De surcroît, à travers son catalogue de banques de données et de variables (voir **1.1.**), le Centre pourra informer la communauté de recherche de la disponibilité de bases déjà constituées dans le cadre de projets antérieurs et donc plus rapidement mobilisables et à plus faible coût, et ce même si les exigences en matière d'accès demeurent les mêmes.

💡 **Intéressant** : [CARTaGENE](#), plateforme publique de recherche localisée au CHU Sainte-Justine, offre une alternative à la destruction des données au terme d'un projet de recherche en permettant aux chercheuses et aux chercheurs de les utiliser à nouveau dans le cadre d'un autre projet de recherche et en soumettant une nouvelle demande.



## 5. Responsabilités du Centre d'accès

Financé par des fonds publics, le Centre d'accès endosse un ensemble de responsabilités en termes de transparence et de reddition de compte envers la population et la communauté de recherche.

### 5.1. Informations sur le fonctionnement et les personnes-ressources

Sur son site Internet, le Centre d'accès devrait présenter clairement son mode de gouvernance, l'organigramme de ses directions et unités, ainsi que les rôles et responsabilités des principaux intervenants et équipes (gestion, analyse, finances, communication, etc.).

Des personnes-ressources devraient être assignées pour assurer la communication avec la communauté de recherche et le public, ainsi que la réception de leurs interrogations.

À ce titre, une Foire aux questions s'avère utile pour répondre non seulement aux questionnements les plus courants des chercheuses et des chercheurs sollicitant des données, mais aussi à ceux des citoyens et citoyennes naviguant en dehors des communautés de recherche.

### 5.2. Protection de la sécurité des données

Le Centre d'accès devrait mettre à disposition de la communauté de recherche et du public l'ensemble des informations ayant trait à la protection de la sécurité et de la confidentialité des données. Ces informations devraient être délivrées dans un langage clair et accessible mettant de l'avant avec précision et concision les principales procédures et processus en place pour garantir la protection des données tout en assurant leur valorisation au service de l'intérêt public.

Le Centre d'accès revêt un rôle central en matière de communication et de pédagogie vis-à-vis du public. Il devrait expliquer à la population en des termes simples et précis les types de données mis à disposition des équipes de recherche (données dépersonnalisées et anonymisées) et les principaux usages qui en sont faits, de façon à mettre en évidence l'équilibre recherché entre la protection des données d'une part et leur valorisation au service de l'intérêt public d'autre part. Ce travail de pédagogie devrait être fait en collaboration étroite avec les équipes de recherche et leurs collaborateurs, qui sont les acteurs clés de la valorisation et de la protection des données.

💡 **Intéressant** : Au Royaume-Uni, le [Health Data Research Innovation Gateway](#), présente à travers des supports numériques innovants (mêlant textes, images et vidéos) les principaux processus ayant trait à la protection des données transitant via cette porte d'accès. Une



page de son site Internet est d'ailleurs entièrement dédiée à l'information des patients et patientes, et du public.

À voir aussi, un [visuel](#) qui décrit les processus en œuvre pour protéger les données au sein d'un environnement de recherche visant à générer de la confiance ("Trusted Research Environment").

Finalement, le Centre d'accès devrait idéalement mettre en place des mécanismes et processus automatisés de façon à journaliser les demandes et les accès aux données faits en son sein ou à distance. Ceci participe de sa bonne reddition de compte auprès des instances réglementaires et du public.

### 5.3. Indicateurs de performance

En vertu des principes de transparence, d'efficacité et d'imputabilité, le Centre d'accès devrait produire et diffuser de façon régulière un rapport présentant un certain nombre d'indicateurs de performance en lien avec le traitement des demandes d'accès à des fins de recherche. Ceci lui permet de s'inscrire dans une démarche d'amélioration continue en partenariat avec les communautés de recherche qui peuvent agir comme des contributrices et collaboratrices de cette démarche.

Les indicateurs de performance du Centre comprennent minimalement :

- Le nombre de demandes d'accès déposées par les chercheuses et chercheurs ;
- Le nombre de demandes acceptées et refusées ;
- Les délais moyens d'accès aux données pour chacun des types de processus d'accès (tels que décrits en en **1.2.**) ;
- Le niveau de satisfaction des chercheuses et chercheurs (mesuré à travers un questionnaire standardisé) ;
- Le nombre de publications/infographies/vidéos générées par les équipes de recherche à partir de la valorisation des données du Centre.

A ces indicateurs de performance minimaux peuvent s'ajouter un certain nombre d'informations complémentaires qui sont d'intérêt pour la communauté de recherche, telles que le nombre de nouvelles bases de données créées par le Centre chaque année, le nombre de demandes d'accès inter-institutionnelles (lorsque plusieurs centres sont sollicités) ou encore le nombre de jumelages ou appariements de données entre institutions.

**Point important** : En matière de délais, le Centre d'accès devrait s'engager à respecter un délai maximal de traitement des demandes d'accès n'excédant jamais 3 mois.



#### 5.4. Registre des projets de recherche

Finalement, le Centre d'accès devrait diffuser sur son site Internet un registre complet et à jour des projets de recherche passés, en cours et achevés étant basés sur l'utilisation des données dont le Centre est le fiduciaire. Ce registre devrait comprendre un ensemble d'informations clés, telles que :

- Le titre du projet de recherche ;
- Le nom de la chercheuse ou du chercheur principal, et son affiliation ;
- La date d'approbation du projet ;
- Le résumé exécutif du projet indiquant sa méthode et ses objectifs en des termes simples et concis ;
- Les jeux de données sur lesquels se base le projet (référençables par un DOI et documentés via un catalogue tel que OBiBa Mica par exemple) ;
- Si le projet est achevé, un lien vers les publications issues de celui-ci ainsi que les autres matériaux de valorisation des résultats de la recherche : balados-diffusion, vidéos de conférence, infographies, etc.

**💡 Intéressant** : Plusieurs plateformes d'accès au Canada et dans le monde établissent un registre à jour des projets de recherche utilisant leurs données. Vous pouvez voir des cas du [Réseau canadien des Centres de données de recherche](#), du [Health Data Hub en France](#) ou du [CPRD au Royaume-Uni](#).