Table nationale des directeurs de la recherche (TNDR)

Principes directeurs pour assurer le fonctionnement et la gestion optimale d'un centre d'accès aux données de santé

Document préparé par le sous-comité
Gouvernance et cadre de gestion

Version 2.0
14 octobre 2021

Les **« Principes directeurs pour assurer le fonctionnement et la gestion optimale d'un centre d'accès aux données de santé »** (dénommés ci-après « Principes directeurs ») ont été publiés en novembre 2020, puis communiqués aux établissements du réseau de la santé et des services sociaux et aux établissements universitaires et de recherche.

Bâtis sur le mode de l'échange et de la collaboration interinstitutionnelle, les Principes directeurs sont le fruit du travail du sous-groupe « Gouvernance et cadre de gestion » de la Table nationale des directeurs de la recherche (TNDR), dont les membres sont déclinés ci-dessous :

Membres du sous-groupe « Gouvernance et cadre de gestion »

Carole Artault-Noury, Cheffe de secteur, Secteur recherche et gestion de la recherche, Direction des technologies de l'information, Université Laval

Geneviève Cardinal, Chef du Bureau de l'éthique de la recherche, CHU Sainte-Justine

Patricia Caris, Directrice générale, Direction générale aux statistiques et à l'analyse sociales, Institut de la statistique du Québec

Camille Craig, Adjointe de direction, Centre de recherche du Centre hospitalier de l'Université de Montréal

Mylène Deschênes, Directrice des Affaires Éthiques et Juridiques – Fonds de recherche du Québec

Marie-Ève Doucet, Conseillère en coordination de la recherche, Direction de la recherche et de la coordination interne, Ministère de la Santé et des Services Sociaux

Marie Hirtle, Adjointe à la direction, Direction de la qualité, de l'évaluation, de la performance et de l'éthique et présidente du CER, Centre universitaire de santé McGill

Carole Jabet, Directrice Scientifique – Fonds de recherche du Québec Santé

Le sous-comité remercie les autres individus ayant contribué à l'élaboration du présent document.

Dans une logique d'adaptation et d'amélioration continues, les **Principes directeurs de la TNDR** sont amenés à être enrichis et bonifiés de façon à intégrer non seulement les évolutions législatives et réglementaires, mais aussi les nouvelles normes et bonnes pratiques en matière de gouvernance et de gestion des données en santé et services sociaux.

De la sorte, à l'automne 2021, un nouveau chantier de travail est lancé à la TNDR, de manière à produire une nouvelle version des Principes directeurs intégrant notamment :

- Les nouvelles normes en matière de cybersécurité et de confidentialité des données de santé
- Les recommandations en matière d'interopérabilité et de valorisation des données de santé inter-établissements
- Les nouvelles propositions de gouvernance des centres d'accès aux données s'inscrivant dans une logique d'intégration des différentes directions et compétences

- Les meilleures pratiques en matière d'engagement des patients, professionnels et citoyens pour une gestion inclusive et collaborative des données de santé

Pour citer ce document : Table nationale des directeurs de recherche (TNDR), Sous-groupe Gouvernance et cadre de gestion. *Principes directeurs pour assurer le fonctionnement et la gestion optimale d'un centre d'accès aux données de santé*; 2020-11-09. 12 p.

Principes directeurs pour assurer le fonctionnement et la gestion optimale d'un centre d'accès aux données de santé

I. Table des matières

II. Sigles et acronymes

CADS :	Centre d’accès aux données en santé
CAI :	Commission d’accès à l’information
CER :	Comité d’éthique de la recherche
DSP :	Direction des services professionnels
DOI :	Digital Object Identifier

III. Définitions

Base de données : Collection de données structurées pour permettre des opérations, parfois très complexes, de lecture, de suppression, de déplacement, de tri, de comparaison ou autres opérations. Lorsque plusieurs bases de données sont constituées sous forme de collection, on parle alors d'une banque de données.

Fiduciaire de la donnée : Correspond à la personne physique ou morale qui est responsable de la planification et de l'élaboration des politiques en lien avec la gestion de la donnée (notamment en ce qui concerne son accès et son utilisation). Le fiduciaire de la donnée applique les exigences légales et règlementaires en lien avec la donnée et supervise la mise en place de politiques et de processus de gestion des données. À ce titre, le fiduciaire de la donnée assure notamment les orientations stratégiques et financières liées à la donnée.

Intendant de la donnée : Correspond à la personne physique ou morale qui est directement responsable de la gestion de la donnée au niveau opérationnel. L'intendant de la donnée gère et maintient la donnée dont il a la charge et à ce titre, assure la qualité, la sécurité et la confidentialité de ladite donnée. L'intendant de la donnée est notamment responsable d'implanter et de gérer l'application de toutes les politiques liées à la qualité de la donnée, la standardisation de la donnée et l'accès à la donnée.

Mandataire de la donnée : Correspond à la personne physique ou morale qui est mandaté pour accomplir le rôle d'intentant de la donnée

Utilisateur de la donnée : Correspond à la personne physique ou morale qui a été autorisée à accéder à la donnée afin d'accomplir une tâche donnée ou dans l'exercice des rôles ou fonctions qui lui sont attribués. L'utilisateur de la donnée est responsable de suivre les lois, politiques, procédures et standards associés à la donnée qu'il utilise et d'utiliser cette dernière uniquement aux fins auxquelles il a été autorisé. L'utilisateur de la donnée a également la responsabilité de signaler à l'intendant de la donnée tout accès non autorisé, utilisation abusive ou problème de qualité de la donnée.

Centre d'accès aux données de santé (CADS) : Infrastructure sécurisée où des données confidentielles relatives aux soins et services de santé et/ou des données constituées à des fins de recherche (p. ex. cohortes de patients ou populationnelles) ou de surveillance sont déposées, manipulées, analysées en suivant des règles prédéfinies d'accès. Le CADS offre un ensemble de services contrôlés de bout en bout pour garantir un haut niveau de sécurité, en soutien à une utilisation secondaire des données à des fins de recherche.

Dépersonnalisation (aussi nommé dénominalisation ou pseudo-anonymisation) : Procédure qui consiste à remplacer les informations nominatives contenues dans un

document par un code d'identification, de manière à empêcher l'identification des individus auprès desquels elles ont été recueillies[1].

Anonymisation : Procédure qui consiste à remplacer les informations nominatives contenues dans un document par un code d'identification, avant de les supprimer définitivement, de manière à rendre virtuellement impossible l'identification des individus auprès desquels elles ont été recueillies[2].

Ré-identification : Tout processus rétablissant le lien entre l'information et l'identité d'un individu.

Renseignements personnels[3] : Renseignements portant sur un individu et permettant d'établir son identité.

Appariement de données : Opération consistant à jumeler des données sur la base de de renseignements identificatoires ou dépersonnalisés à partir d'au moins deux bases de données différentes. Il s'agit donc de combiner de l'information provenant de deux bases de données différentes pour les mêmes individus.

Mise en commun de données sans appariement : Opération consistant à combiner des informations de deux bases de données différentes provenant d'individus différents.

Variables avec identifiants directs : Une ou plusieurs variables utilisées seules ou en combinaison permettant l'identification un individu en particulier.

Variables avec identifiants indirects : Variables utilisées seules ou en combinaison permettant d'identifier un individu en ayant connaissance du contexte.

Variable non identificatoire : Variable ne permettant pas la ré-identification d'un individu.

---

[1] Office québécois de la langue française : http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26532656.

[2] Office québécois de la langue française : http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26532654.

[3] Office québécois de la langue française : http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8398805.

Principes directeurs pour assurer le fonctionnement et la gestion optimale d'un centre d'accès aux données de santé

1. Introduction

Dès 2017, la Table Nationale des directeurs de la recherche du Ministère de la Santé et des Services sociaux (TNDR) a défini comme priorité l'organisation et l'accès aux données clinico-administratives et de recherche. Celles-ci s'avèrent essentielles tant au développement de technologies qui facilitent et améliorent la détection et les diagnosti cs de maladies, qu'à l'accélération de la découverte de nouveaux traitements, de la compétitivité en recherche clinique ou encore de l'optimisation de l'utilisation des ressources dans l'ensemble du système de santé. Cette priorité a gagné en importance avec l'accélération du développement de l'intelligence artificielle appliquée à la santé. De la médecine personnalisée aux modèles prédictifs, en passant par le dépistage précoce d'effets secondaires, l'intégration du numérique et de l'intelligence artificielle dans les établissements de soins et leurs centres de recherche est en train de transformer la pratique clinique.

Les membres de la TNDR ont traduit leur volonté dans une vision prioritaire qui guide leurs actions : accroître de façon majeure la capacité de recherche et la génération de nouvelles connaissances en organisant et facilitant le traitement des données de santé.

Organiser et faciliter le traitement des données en santé comprend plusieurs niveaux synthétisés comme suit :

- Évaluer, mettre en œuvre, rendre interopérables des entrepôts de données cliniques;
- Mutualiser les compétences, les expertises et les outils pour exploiter les données numériques;
- Se doter d'un cadre organisationnel pour assurer une utilisation responsable des données.

Le groupe de travail *Gouvernance et cadre de gestion* a été mis en place pour travailler spécifiquement sur les cadres organisationnels. À partir des meilleures pratiques en cours (voir Annexe 1 pour les documents consultés), ce groupe de travail propose dans ce document, une liste de principes directeurs qui devraient guider l'élaboration et l'opérationnalisation des cadres organisationnels (aussi nommés cadres de gestion) des données en santé.

Il est important de souligner ici que ces principes s'inscrivent dans le cadre des **principes FAIR**, élaborés des 2016[4] pour servir de principes directeurs concis et mesurables visant à garantir la découvrabilité (findable), l'accessibilité, l'interopérabilité et la réutilisation des (méta)données de recherche pour les humains et les machines. Les principes FAIR sont devenus une norme quant à la qualité des services de gestion des données de recherche et sont largement adoptés par les différents acteurs (chercheurs, organisations de recherche, agences subventionnaires, etc.). Bien que non abordés spécifiquement dans le document, les principes doivent être cohérents avec le concept de science ouverte (open science) et le partage d'outils en code source ouvert (open source).

Enfin, les principes directeurs décrits ci-après sont un guide pour les organisations qui souhaitent mettre en place et opérer un centre d'accès aux données de santé (CADS). Ils viennent en appui aux cadres et processus qui prévalent dans les établissements, tels que le cadre de gestion de l'établissement, cadres de référence des ministères et organisations tutelles, etc. Bien que fortement conseillé, il appartient aux organisations de s'y référer.

2. Les principes directeurs pour le fonctionnement et la gestion optimale d'un CADS

Chaque établissement qui désire mettre en place un CADS doit se doter d'un cadre de gestion balisant l'organisation, la gestion et l'utilisation des données, cadre qui doit ensuite se décliner en politiques et procédures qui en permettront son opérationnalisation. Les éléments ci-après devraient être décrits, d'une façon ou d'une autre, dans le cadre de gestion.

2.1 Description générale du CADS et des données disponibles

Une description générale du CADS incluant ses principaux objectifs doit d'abord être présentée. La description doit mettre en évidence l'institution d'attache et dans quels cadres légaux et réglementaires, le CADS peut opérer.

Cette section inclut également une description des données qui sont rendues accessibles. Les données peuvent être décrites en spécifiant la nature et la source des données (p. ex. données clinico-administratives, données d'enquêtes colligées à des fins de recherche, systèmes d'informations et entrepôts de l'établissement, etc.). La durée de conservation des données devrait également être précisée.

Les voies de valorisation des données sont à concevoir dès la constitution du CADS (p. ex. utilisation secondaire à des fins de recherche, utilisation à des fins d'analyse de qualité-

[4] Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., … Mons, B. (2016). *The FAIR Guiding Principles for scientific data management and stewardship*. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

performance, utilisation pour des tableaux de bord, fins de surveillance etc.). Il est également souhaitable de définir, dès la mise en place du CADS, les utilisateurs éligibles à l'accès aux CADS (p. ex. chercheurs académiques, membres de l'industrie, etc.) et de définir les conditions d'accès de chacun (voir section *Gestion des accès*). Les mécanismes de valorisation généraux des résultats de recherche CADS doivent également être réfléchis (voir section *Valorisation des résultats issus du CADS*).

Enfin, il est impératif d'énoncer de façon claire les responsabilités du CADS et des tiers vis-à-vis des données accessibles dans le centre : qui en est le fiduciaire, qui en est l'intendant ou le mandataire (dans le cas d'une délégation de gestion de données initialement gérées par un tiers).

2.2 Gouvernance

La structure de gouvernance et des différentes autorités assurant le fonctionnement et la gestion du CADS doit être définie et idéalement schématisée sous la forme d'un organigramme.

Les mandats des différentes entités constituant la gouvernance doivent s'assurer de couvrir les points de rôles et responsabilités décrits ci-après. La personne en autorité de l'établissement où les données seront hébergées et gérées est responsable vis-à-vis du non-respect de ces rôles, responsabilités et obligations.

Gouvernance et orientations stratégiques du CADS :

- Adopter réviser et assurer la mise à jour du cadre de gestion du CADS;
- Produire et assurer la mise en place des politiques, procédures, meilleures pratiques et ressources pour une gestion et une utilisation conforme des données, incluant la procédure de gestion d'un bris de confidentialité;
- Assurer la mise en œuvre des mécanismes de suivis;
- Adopter et suivre le plan stratégique et le plan d'actions du CADS.

Contrôle de la qualité et de la conformité :

- S'assurer de l'intégrité des données hébergées tout au long du cycle de vie de la donnée (entrée des données, intégration des données ou combinaison des données de différentes sources, préparation et extraction des données, analyse, conservation et destruction) par la mise en place de mécanisme de suivi (p. ex. production régulière de rapport de validation et de mesure afin de détecter des incohérences potentielles avec les données);
- Réaliser les audits internes de conformité aux procédures préétablies et mettre en place les changements requis.

Orientations scientifiques et priorités de développement du CADS :

- Proposer des orientations et priorités de développement du CADS;
- Réaliser une veille (benchmarking) des projets similaires au CADS et identifier les opportunités de collaborations porteuses.

Dépendamment de la mission du CADS (plateforme de services et/ou partenaire de développement de recherche) :

- Planifier la programmation de recherche du CADS (établissement des critères d'évaluation des projets et leur évaluation);
- Prioriser et intégrer les nouveaux projets à la programmation en cours;
- Suivre et effectuer les bilans des projets;
- Évaluer les retombées scientifiques des projets;
- Créer de nouvelles occasions de recherche en favorisant l'interdisciplinarité.

Gestion opérationnelle et administrative du CADS :

- Mettre en place et diffuser le cadre de gestion ;
- Suivre les standards de qualité et de rigueur scientifique ainsi que les politiques et procédures et meilleures pratiques mises en place par le cadre de gestion établi;
- Développer et implanter des méthodes et des outils;
- Gérer et traiter les demandes d'accès aux données (évaluation de l'éligibilité et de la faisabilité des demandes ainsi que le suivi des projets d'accès aux données autorisés);
- Procéder à l'intégration des systèmes d'information, à l'extraction et préparation de même qu'à l'analyse des données (tout dépendant des types de projets d'accès aux données).

2.3 Confidentialité

Le CADS doit s'assurer de la confidentialité des données sous sa gestion et la protection des renseignements personnels associés. Le risque lié à la confidentialité doit ainsi être géré adéquatement tout au long du cycle de vie de la donnée, le CADS devant être en mesure de maintenir un niveau de risque acceptable tout en étant conscient que le risque ZÉRO n'existe pas. Afin d'y parvenir, un CADS doit s'assurer de respecter les principes suivants :

- Une séparation des données identificatoires (c.-à-d. permettant d'identifier les individus) des autres données et la création d'une liaison via un identifiant unique doivent être réalisées. Le principe de séparation consiste à conserver les données contenant les identifiants dans un espace distinct et d'en limiter l'accès aux individus autorisés. Seuls les responsables de la base de données détiennent des privilèges d'accès aux systèmes contenant les identifiants ainsi qu'aux autres données détenues par le CADS pour des fins de développement, d'intégration et d'opérationnalisation, d'exploitation et d'analyse.
- Les chercheurs utilisateurs auront seulement accès aux données sous forme dépersonnalisées pour fins d'analyse. Il est recommandé de créer des identifiants uniques distincts pour chaque projet d'analyse ce qui permet de contrer la possibilité d'échanges d'informations non autorisés entres différents utilisateurs.
- Le CADS doit se doter d'un mécanisme d'évaluation du risque lié à la protection des données personnelles, permettant d'attribuer à un sous-ensemble de données d'analyse

un niveau de risque quant à la ré-identification. Il est recommandé que les différentes variables produites soient classées en catégories distinctes (i- variables avec identifiants directs, ii- variables avec identifiants indirects et iii- variable non identificatoire) ce qui facilite l'attribution des niveaux de risque. Les niveaux de risque peuvent être catégorisés de la façon suivante :

  - **Risque élevé** : Le sous-ensemble de données d'analyse comprend des éléments avec identifiants directs ou avec suffisamment d'identifiants indirects pouvant être utilisés pour identifier un individu.
  - **Risque modéré**: Le sous-ensemble de données d'analyse comprend majoritairement des données dépersonnalisées pouvant être partiellement exposés à la ré-identification.
  - **Risque faible** : Le sous-ensemble de données d'analyse contient des données dépersonnalisées (ne contient aucun élément avec identifiants directs et les éléments avec identifiants indirects ont été manipulés pour assurer un niveau acceptable de risque à la ré-identification).

- En se basant sur les niveaux de risque attribué, le CADS doit définir ce qui est permis en matière de la localisation des données à des fins d'analyse ainsi que des conditions d'analyse (voir Annexe 2). Le CADS doit aussi prévoir quelles actions ou manipulations de l'ensemble de données d'analyse seront réalisées pour minimiser le risque associé et le rendre acceptable. De plus, une validation du risque à la ré-identification pour les sous-ensembles de données d'analyse produit et des résultats doit être réalisée par les employés du CADS afin de prévenir la divulgation (fortuite) de renseignements personnels.
- Un nombre limité de personnes qualifiées qui auront complété avec succès une enquête de sécurité ainsi qu'une formation sur la protection de la vie privée et sur la confidentialité pourra manipuler et accéder aux données identificatoires. Ces employés devront signer un engagement de confidentialité avec l'établissement qui a le mandat de gérer le CADS. Un renouvellement d'engagement sur une base régulière (idéalement annuelle) est à prévoir.
- Chaque chercheur utilisateur du CADS est soumis à la conduite responsable en recherche de son établissement ainsi qu'aux règles définies par les organismes subventionnaires ce qui permet d'assurer une éthique de travail et de recherche exemplaire. L'utilisateur et son établissement d'affiliation devront signer un engagement de respect de confidentialité et devront traiter les données d'analyse comme étant de l'information confidentielle. Le non-respect de la confidentialité par un chercheur utilisateur résultera en une suspension immédiate de l'utilisation du CADS et s'accompagnera potentiellement d'une sanction ou d'une pénalité telle que stipulée dans les politiques de bonnes conduites en recherche établissement d'attache du CADS. Le chercheur utilisateur devra s'assurer de respecter l'énoncé de politique de l'organisme subventionnaire en ce qui concerne l'utilisation responsable des données. Se référer à la section *Formation* pour plus de détails sur les formations recommandées.

2.4 Sécurité de l'information

La sécurité de l'information et la protection des renseignements personnels doit être assurée tout au long du cycle de vie des données. Les mesures pour s'en assurer devraient être intégrées dès la mise en œuvre d'un projet impliquant un traitement de données selon l'approche de la protection de la vie privée dès la conception (ou Privacy by Design)[5].

Sont décrites dans cette section les principales mesures de sécurité qui se distinguent en trois principaux types : les mesures physiques, les mesures techniques et les mesures organisationnelles.

- *Mesures physiques* : Ces mesures concernent tout ce qui est en lien avec la sécurité physique des données. Par exemple, les serveurs sur lesquels les données sont entreposées se situent dans un local verrouillé, la conservation adéquate comprenant différentes copies et sauvegardes et permettant d'assurer une redondance de l'infrastructure, la protection de l'aire de travail où les analyses sont réalisées par une entrée par carte d'accès, etc.
- *Mesures techniques* : Ces mesures concernent tout ce qui a trait aux mesures techniques ou technologiques permettant d'assurer la sécurité des données. Par exemple, des serveurs protégés par un pare-feu empêchant les connexions externes, des transferts de données effectués avec des données encryptées et un protocole sécurisé, un registre (log) des connexions généré de façon régulière (traçabilité des accès), etc. Peut aussi être comprise dans cette catégorie toute utilisation de technologies permettant d'extraire et d'analyser des données sans accès à la donnée réelle tout en assurant un résultat scientifiquement valide.
- *Mesures organisationnelles* : Ces mesures concernent l'identification et l'organisation claire des rôles qui permettent d'assurer la sécurité des données. Par exemple, la gestion des permissions de différents individus ou groupes d'individus, les formations et autres engagements à respecter par les membres de l'équipe de gestion des données ou l'utilisateur (voir section *Formation* pour plus de détails). Sont aussi comprises dans ce type de mesures toutes les procédures organisationnelles mises en place permettant d'assurer la sécurité dans les différentes étapes du cycle de gestion des données.

2.5 Conformité

L'unité de contrôle de qualité et de conformité du CADS est responsable de réaliser un suivi régulier pour s'assurer de la conformité au cadre de gestion du CADS et des politiques et procédures s'y rattachant, et ce, garantissant le respect des meilleures pratiques de gestion et d'utilisation des données. Il est recommandé que l'unité de contrôle de qualité

[5] A. Cavoukian. *Privacy by Design, the 7 Foundational Principles. Implementation and Mapping of Fair Information Practice*s. http://dataprotection.industries/wp-content/uploads/2017/10/privacy-by-design.pdf.

réalise des audits internes de conformité et de contrôle qualité des procédures sur une base annuelle. En cas de non-conformité, des mesures doivent être prévues au cadre de gestion et aux procédures afin d'apporter les corrections aux processus et aspects identifiés comme problématiques. Le résultat de l'audit annuel doit être transmis au comité de gouvernance du CADS.

Enfin, il est fortement suggéré que le CADS se soumette, à intervalle régulier ou suite à tout événement qui modifierait le cadre légal ou réglementaire, à un audit par une firme externe pour tout ce qui a trait au respect des lois et règlements la gouvernance, la protection des renseignements personnels et la sécurité des données.

2.6 Formation

En matière de formation, il est essentiel que l'équipe de gestion opérationnelle du CADS définisse, avant la mise en exploitation des données, les formations qui seront requises par les employés du CADS et par les chercheurs utilisateurs. Le CADS est ainsi responsable d'identifier le contenu de sa formation et ses mises à niveau en plus d'offrir la formation (un rappel annuel des formations est fortement suggéré). Le contenu de la formation offerte doit être en cohérence avec les spécificités propres à son établissement d'attache ce qui permet de bonifier les formations qui y sont déjà établies et devant déjà être suivies (les duplications sont à éviter). Dans le guide de formation développé, les formations suivantes sont suggérées :

Pour les employés du CADS :

- Toute formation exigée par l'établissement d'attache (p. ex. : conduite responsable en recherche, formation de base en éthique de la recherche, modes opératoires normalisés, etc.)
- Formation en matière de sécurité et confidentialité des données
- Formation sur le cadre de gestion du CADS ainsi que les différentes politiques et procédures s'y rattachant
- Familiarisation avec les formations données aux chercheurs-utilisateurs

Pour les chercheurs utilisateurs :

- Formation sur les politiques et procédures pour l'utilisation des données du CADS
- Formation sur les meilleures pratiques de confidentialité, sécurité et gestion des données
- Formation sur les politiques de bonnes conduites en recherche de l'établissement d'attache du CADS.

Il appartient à chaque CADS de définir les modalités d'implantation de ses formations, celles-ci devant être définies dans le contexte spécifique du CADS (p. ex. : format offert pour les formations - présentiel, - en ligne, - capsule vidéos, mécanisme pour suivre les formations réalisées et mettre en place les rappels, etc.).

2.7 Gestion des accès aux données et transparence

Afin de permettre un fonctionnement optimal et assurer la transparence du processus d'accès aux données, le CADS devrait s'engager à :

- Publier le processus d'accès aux données, incluant les prérequis nécessaires pour obtenir un accès, incluant les délais de traitement ciblés des demandes d'accès;
- Rendre publique une liste des projets avec titre et résumé vulgarisé utilisant les données du CADS, le nom de l'utilisateur et son organisation d'attache;
- Publier une grille tarifaire des services offerts. À noter que les tarifs peuvent être variables selon la nature de l'utilisateur (public ou privé);
- Enfin, il apparaît essentiel que le CADS définisse les conditions d'accès qui devront s'appliquer dans un contexte de crise sanitaire, les besoins en données (accès, partage et particulièrement en jumelage de données) étant particulièrement importants dans ce contexte et caractérisés par leur nature urgente. Il est recommandé que le CADS détaille les modalités de fonctionnement dans ce contexte particulier afin d'anticiper et mieux prévoir comment il pourra répondre aux demandes dans un délai raisonnable tout en assurant le respect des meilleures pratiques.

Les Termes et Conditions de gestion et d'accès aux données du CADS sont habituellement définis dans une **Entente d'accès (modèle pré-établi)** entre l'établissement d'attache du CADS et l'utilisateur une fois que l'utilisateur a su démontrer avoir tous les prérequis et obtenu toutes les autorisations nécessaires (ex. CER, CAI DSP, autres) pour les fins de son projet d'analyse. L'entente d'accès doit clairement indiquer une date d'entrée en vigueur et une date de fin. L'entente d'accès devrait couvrir au minimum les points suivants :

- Responsabilités et obligations de chacune des parties
- Transfert des connaissances, diffusion des résultats et reconnaissance des auteurs : Toute publication ou présentation utilisant les données du CADS devrait inclure une mention reconnaissant l'apport du CADS. L'utilisateur doit également s'engager à reconnaître la contribution d'un membre de l'équipe du CADS lorsque celle-ci respecte les règles et meilleures pratiques en matière de reconnaissance d'auteurs sur les publications.
- Période d'usage et de conservation des données utilisées : il s'agit dans cette section de prévoir le devenir de l'ensemble de données confié à l'utilisateur, lorsque le projet de recherche est terminé. La conservation des données est encouragée à condition de pouvoir en démontrer la pertinence et que les meilleurs standards de sécurité sont rencontrés. Si les garanties ne sont pas adéquates, une date et un processus de destruction des données doivent être proposés. À noter que tout règlement décrivant des conditions de conservation de données de santé et de données administratives doit être respecté.
- Découverte d'une erreur fortuite dans le sous-ensemble des données auquel l'utilisateur a obtenu accès : L'utilisateur doit aviser les responsables du CADS

immédiatement afin que puissent se déployer les interventions nécessaires pour corriger l'erreur le plus rapidement possible.

- Découverte d'une information permettant d'identifier un individu ou bris de confidentialité : L'utilisateur doit aviser les responsables du CADS immédiatement afin que puissent se déployer les interventions nécessaires pour remédier à la situation le plus rapidement possible.

2.8 Valorisation des résultats issus du CADS

Les éléments suivants sont à considérer pour la valorisation des résultats issus du CADS :

- Propriété intellectuelle : Dans aucun cas l'utilisateur a des droits de propriété intellectuelle sur les données extraites de la base de données, ni aucun droit de commercialiser les données extraites. De plus, l'entente d'accès aux données mise en place entre l'établissement d'attache du CADS et l'utilisateur de la base de données doit préciser les termes et conditions relatifs à la propriété intellectuelle et la propriété et la commercialisation des résultats doivent être clairs et précis. Les termes de l'entente doivent respecter les politiques institutionnelles de l'établissement d'attache du CADS en ce qui concerne la propriété intellectuelle, la commercialisation des résultats et la propriété des résultats.
- Entente de collaboration avec entité privée : Dans le cas d'un projet collaboratif entre un chercheur utilisateur et une entité privée, une entente de collaboration entre l'entité privée et l'établissement doit d'abord être mise en place. Cette entente de collaboration doit inclure les articles retrouvés dans l'entente d'accès aux données. Il est possible qu'une entente de collaboration soit élaborée dans un premier temps et qu'une entente d'accès aux données soit définie dans un deuxième temps.

Enfin, la mise en place d'un CADS doit de faire dans l'optique que celui-ci se bonifiera au fur et à mesure que de nouvelles données seront intégrées et que de nouvelles collaborations amèneront des processus facilités d'arrimage entre plusieurs sources de données augmentant ainsi les capacités générales de recherche. Certains CADS pourraient même définir que, pour certains types de projets, les données issues de l'analyse (ou résultats) soient redéposées à même l'infrastructure du CADS afin de venir enrichir les données rendues disponibles par celui-ci.

Dans cette optique, il est essentiel pour le CADS de se pencher sur les mécanismes de citation de sa ressource qui lui permettront d'assurer tout d'abord i) une reconnaissance de l'utilisation de son infrastructure dans les projets des chercheurs utilisateurs, et ii) un meilleur suivi des indicateurs d'impacts de la recherche du CADS. À titre d'exemples, la publication d'un article décrivant l'infrastructure du CADS et les données disponibles, l'enregistrement de la ressource dans différents catalogues recensant les infrastructures de données ou l'obtention d'un numéro de citation unique (ex. obtention d'un DOI pour un ensemble spécifique défini de données de CADS) peuvent être des options intéressantes à envisager pour certains CADS.

2.9 Appariement et mise en commun des données

Il est à prévoir que l'utilisateur des données du CADS souhaite mettre en commun ou apparier les données obtenues du CADS avec les données d'autres établissements ou d'autres sources de données pour réaliser son projet d'analyse. Cette mise en commun ou appariement des données doit être rendue possible par le CADS, dans le respect des lois et règlements en vigueur, notamment mais pas seulement pour la protection des renseignements personnels.

- L'utilisateur doit se conformer et avoir obtenu toutes les approbations nécessaires pour mettre en commun, voire apparier, les données du CADS avec d'autres données pertinentes pour accomplir son analyse.
- Le CADS peut mettre à disposition des utilisateurs une grille détaillant toutes les autorisations nécessaires pour mettre en commun et apparier des données.
- L'obtention des autorisations est la responsabilité de l'utilisateur du CADS; il est cependant de la responsabilité du CADS de vérifier que toutes les approbations nécessaires ont bien été obtenues avant d'autoriser une mise en commun ou un appariement de données.
- Un appariement de données exposant des renseignements personnels devra absolument être effectué par le personnel autorisé et formé du CADS et non l'utilisateur des données (conformément aux règles de sécurité).

2.10 Adoption, implantation et révision du cadre de gestion

Cette section du cadre de gestion est importante pour rappeler aux responsables d'un CADS qu'un cadre de gestion est un document dynamique, qui doit évoluer pour intégrer les meilleures pratiques les plus à jour et répondre ainsi à un mécanisme d'amélioration continue tant pour l'organisation, que la gestion et l'utilisation des données en santé. Parmi les points qui doivent être répondus dans cette section :

- Définir qui a la responsabilité de l'adoption, implantation et révision du cadre de gestion du CADS. En général, le comité en charge de la gouvernance du CADS est responsable en tout temps de l'adoption du cadre de gestion et de sa révision et de mettre en place les mécanismes pour assurer la conformité des politiques et procédures s'y rattachant. Les responsables des opérations assurent l'implantation et l'application du cadre.
- Définir la fréquence de révision du cadre de gestion : au moins tous les deux ans ou selon tout événement qui modifierait le cadre légal ou réglementaire sur quel le cadre de gestion s'appuie pour s'assurer de sa cohérence avec la législation actuelle et des meilleures pratiques. Une documentation des procédures découlant du cadre de gestion de type Modes opératoires normalisés (MON) devrait également être mise en place.

3. Annexes

3.1 Annexe 1. Principaux documents et procédures analysés pour établir les principes directeurs

Centre de recherche du CHUM, *Cadre de gestion du Centre d'intégration et d'analyse des données médicales du CHUM (CITADEL) – volet recherche*; 20 mars 2019. 25 p. Accessible en ligne : https://www.chumontreal.qc.ca/sites/default/files/inline-files/citadel_cadre_gestion_recherche_v1_0.pdf.

Institut de la statistique du Québec, *Guichet d'accès aux données de recherche, modèles d'accès aux données de recherche*; Avril 2020. Présentation disponible sur demande.

McGill University Health Center, *MUHC Anonymized Data Warehouse Management Framework*; December 21 2018. 42 p.

PULSAR, Université Laval, *Cadre de gestion Banque de données sur la santé durable*; 20 mai 2019. 63 p. Accessible en ligne : https://pulsar.ca/sites/default/files/inline-files/Cadre_gestion_PULSAR_2018-142CG_20-05-2019_4.pdf?lang=fr.

3.2 Annexe 2. Risques de ré-identification et considérations liées à l'hébergement, au transfert, à l'analyse des données et à la diffusion des résultats

Risque de ré-identification	Hébergement des données d'analyse et transfert	Analyse des données et diffusion des résultats
Risque élevé	• Sous-ensemble de données d'analyse hébergé sur les serveurs du CADS • Aucune possibilité de transfert du sous-ensemble de données	• Analyse réalisée par un employé du CADS • Partage des résultats à l'utilisateur seulement sous forme de données agrégées
Risque modéré	• Sous-ensemble de données d'analyse hébergé sur les serveurs du CADS • Aucune possibilité de transfert du sous-ensemble de données	• Analyse pouvant être réalisée par un utilisateur avec les outils d'analyse des systèmes du CADS • Partage des résultats à l'utilisateur seulement après vérification de l'absence de risque de ré-identification par employé du CADS
Risque faible	• Sous-ensemble de données d'analyse pouvant être hébergé hors des serveurs du CADS • Transfert du sous-ensemble possible par un mode de transfert autorisé par le CADS	• Analyse pouvant être réalisée par un utilisateur avec les outils de son choix • Aucune restriction quant au partage des résultats

Notes. Les mesures de sécurité mises en place pour l'hébergement, le transfert et l'analyse des données de même que pour la diffusion des résultats d'analyse sont dépendantes du risque de ré-identification associé au sous-ensemble de données d'analyse. L'analyse du risque lié aux sous-ensembles de données d'analyse est réalisée un employé du CADS.