Centre de recherche
CHU Sainte-Justine
Le centre hospitalier universitaire mère-enfant

Univers informationnel (UniC)

Cadre de gouvernance pour la création et l'exploitation d'un lac de données

Version 1

Approuvé le 26 mai, 2022

1. Introduction

1.1 Le CHU Sainte-Justine

Mission, vision

Le Centre hospitalier universitaire Sainte-Justine (CHUSJ) est le seul établissement de santé dédié exclusivement aux enfants, aux adolescents et aux mères au Québec. La mission du CHUSJ est d'améliorer la santé des enfants, des adolescents et des mères du Québec, en collaboration avec les partenaires du réseau de santé et ceux des milieux d'enseignement et de recherche.

Le CHUSJ entend assumer pleinement chacun des six mandats découlant de sa mission universitaire :

- Soins spécialisés et ultraspécialisés;
- Recherche fondamentale et clinique en santé de la femme et de l'enfant;
- Enseignement auprès des futurs professionnels de la santé et des intervenants du réseau;
- Promotion de la santé;
- Évaluation des technologies et des modes d'intervention en santé;
- Réadaptation, adaptation et intégration sociale pour les enfants et les adolescents présentant une déficience motrice ou de langage.

Valeurs

Le CHUSJ s'est doté d'un Code d'éthique. Les valeurs et la philosophie qui s'y rattachent, présentées dans ce code d'éthique, proviennent de la vision des patients, de leurs proches et de celle des personnes qui y travaillent.

Ainsi les valeurs phares du code sont les suivantes:

- La quête d'excellence,
- La bienveillance,
- Le partenariat
- L'engagement individuel et collectif

La philosophie du code s'articule autour d'un concept fondamental pour le CHUSJ, « Tous des soignants! ». Ainsi autant les patients, les familles, les intervenants que les gestionnaires ont un rôle fondamental et complémentaire à jouer dans la relation au centre de l'action dans nos murs.

1.2 Le Centre de recherche

Le Centre de recherche du CHUSJ est considéré comme une référence pour la recherche mère-enfant au Canada. Il réunit plus de 200 chercheurs, dont plus de 110 chercheurs cliniciens, ainsi que plus de 450 étudiants de cycles supérieurs et post-doctorants. Le Centre de recherche du CHUSJ a démontré son leadership dans plusieurs domaines de recherche. Il poursuit sa mission de faire avancer les connaissances, de développer la santé de précision qui impactera non seulement le diagnostic et la prise en charge des maladies, mais aussi les trajectoires de santé afin de créer un avenir en santé pour les enfants et les femmes de l'ensemble du Québec.

1.3 Description de l'Univers informationnel (UniC)

La valorisation des données, recueillies sur les patients du CHUSJ dans le cadre des soins, pour la recherche et l'amélioration des soins vise à contribuer à l'amélioration des connaissances scientifiques et au bien-être éventuel des populations pédiatriques et maternelles. Une saine exploitation des sources de données cliniques et clinico-administratives représente un prérequis indispensable pour soutenir l'hôpital, les cliniciens et les chercheurs dans leurs missions respectives et leurs valeurs communes pour prévenir les maladies et améliorer les pratiques de soins.

L'Univers informationnel (UnIC) est une initiative structurante du Centre hospitalier universitaire Sainte-Justine dont l'objectif est de créer une plateforme centralisée pour fournir aux chercheurs, aux cliniciens, aux gestionnaires de l'hôpital et aux partenaires, des données cliniques et clinico-administratives complètes, bien documentées, organisées et mises à jour afin de promouvoir et faciliter la recherche, l'évaluation et l'innovation. Cette initiative est centrale pour le développement de la recherche clinique, de l'intelligence d'affaires et de l'intelligence artificielle au CHUSJ et sera la pierre angulaire du nouveau Centre de valorisation des données mère-enfants.

En tant que centre hospitalier universitaire, le CHUSJ désire être un accélérateur pour la recherche basée sur des approches en intelligence artificielle et participer de façon plus

efficiente au processus d'amélioration continue de la qualité des soins et services de première ligne, en visant notamment la valorisation et l'exploitation des données clinico-administratives dans un contexte de recherche. Contraints par l'absence de solution rapide pour accéder à des données de qualité, et par la difficulté de pouvoir transformer et interrelier celles-ci, les chercheurs du CHUSJ demeurent limités dans leur capacité de développer des outils pour la recherche et l'aide à la décision clinique. L'UnIC permettra d'offrir aux utilisateurs une plateforme intégrée permettant de les soutenir dans leurs recherches sur les données massives et de leur offrir des outils informationnels et analytiques pour aborder des problèmes complexes inhérents à la santé. L'infrastructure de l'UnIC permet de mieux protéger la confidentialité de l'information des patients tout en répondant aux besoins des chercheurs d'obtenir des données leur permettant d'examiner et de formuler des recommandations sur des questions complexes liées au système de santé et au bien-être des usagers.

L'UnIC est constitué d'un lac de données qui est alimenté systématiquement par les différents systèmes sources hospitaliers du CHUSJ (Annexe A). Ce lac de données inclut une infrastructure d'archivage des données et une série de solutions logicielles qui permettent d'extraire les données-patients des systèmes sources de l'hôpital et de les intégrer dans l'environnement du lac. Les données sources, qu'elles soient de nature clinique, clinico-administrative ou administrative sont comprises dans ce lac de données, organisées, documentées et rendues accessibles aux chercheurs, aux cliniciens, aux administrateurs de l'hôpital et aux partenaires, ayant obtenu les approbations et autorisations requises.

Le CHUSJ est fiduciaire de l'information détenue sur les patients et intégrée dans le lac de données.

1.4 Portée de ce document

L'UniC est un projet institutionnel du CHUSJ. Il est essentiel que ce projet soit réalisé selon les meilleures pratiques de gouvernance et de façon à mener à bien ses objectifs tout en respectant les normes de sécurité en vigueur et le droit à la vie privée des patients du CHUSJ ainsi que le principe de bienfaisance.

Le respect de la vie privée est une valeur fondamentale qui est essentielle à la protection et à la promotion de la dignité humaine. Le non-respect de la vie privée et de la confidentialité peut causer des préjudices à des personnes ou à des groupes de personnes. Ainsi, les renseignements personnels doivent être recueillis, utilisés et divulgués de manière à respecter le droit à la vie privée des patients et de leur famille.

Le principe collectif de bienfaisance, quant à lui, vise l'amélioration des connaissances pour une population en meilleure santé.

Ce document a pour objectif d'énoncer notamment la structure, les règles et les balises permettant d'encadrer de façon sécuritaire, éthique et efficace l'acquisition, l'accès, l'utilisation et la diffusion des données accessibles via l'UnIC. Ce cadre s'articule autour du principe individuel de protection de la vie privée et du principe collectif de bienfaisance.

Les principes et pratiques de gouvernance rattachées à ce cadre de gouvernance englobent toutes mesures physiques, techniques et organisationnelles permettant d'assurer la sécurité de l'information tout au long de son cycle de vie.

2. Définitions

Accès aux données : Le droit ou la possibilité de consulter et/ou d'utiliser les données conservées dans une base de données ou un dépôt/un environnement de recherche/ un environnement virtuel de recherche. L'accès aux données est un élément important dans la gestion des données de recherche.

Chiffrement : Processus de transformation de l'information en un format différent ou de cryptage des données. Seule la personne possédant la clé de déchiffrement peut lire les données d'origine. Le chiffrement a pour objet d'assurer la confidentialité de données numérisées stockées sur un ordinateur ou un serveur et transmises par internet ou enregistrées sur des périphériques de stockage. Les données chiffrées sont brouillées et illisibles par toute personne qui ne possède pas la clé de déchiffrement, le code secret ou le mot de passe [1].

Couplage : La combinaison ou appariement de deux ou de plusieurs ensembles de données possédant des éléments en commun susceptibles de fournir de nouveaux renseignements ou de nouveaux ensembles de données.

Cycle de vie des données : Tous les stades d'existence de données à partir de la collecte jusqu'à la destruction. La perspective du cycle de vie permet la gestion active de données au fil du temps, pour en assurer ainsi la sécurité, l'accessibilité et l'utilité [1].

Dépersonnalisation (brouillage) : Le fait de modifier les renseignements personnels liés à la personne afin de réduire le risque de divulgation de son identité. Cela peut inclure

le brouillage d'identificateurs directs (p. ex. nom, numéro de téléphone, coordonnées géographiques), la transformation (recodage, combinaison) ou la suppression d'identificateurs indirects qui pourraient être utilisés seuls ou en combinaison pour identifier une personne (p. ex. date d'anniversaire, coordonnées géographiques, dates d'événements clés). Lorsque la dépersonnalisation (brouillage) est exécutée convenablement, le risque de réidentification de données partagées ou publiées est atténué [1].

Détenteur de l'information: Un employé désigné par son organisme public, appartenant à la classe d'emploi de niveau-cadre ou à une classe d'emploi de niveau supérieur, et dont le rôle est, notamment, de s'assurer de la sécurité de l'information et des ressources qui la sous-tendent, relevant de la responsabilité de son unité administrative. Le terme « détenteur de processus d'affaires » est utilisé lorsque ce rôle se limite à un processus d'affaires déterminé [2] .

Données brutes : Données qui n'ont pas été interprétées en vue d'une utilisation. Les données brutes ont le potentiel de se transformer en « information », mais au préalable elles doivent être soumises à une extraction sélective et requièrent organisation, analyse et formatage avant d'être présentées [1].

Données confidentielles : Renseignements à caractère privé, dont l'accès doit être restreint ou contrôlé.

Fiduciaire de la donnée : 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.

Lac de données : Dépôt de stockage qui contient une vaste quantité de données brutes dans leur format d'origine, y compris des données structurées, semi-structurées et non structurées. La structure des données et les intentions d'utilisation ne sont pas définies tant que les données ne sont pas requises.

Codage : Utilisation d'un code pour désigner une personne, un groupe ou un lieu spécifique afin de supprimer tout lien; le lien ne peut plus être établi en l'absence d'un registre de clé qui permet de faire la correspondance entre le nom fictif et la personne, groupe ou lieu spécifique. Dans le domaine de la recherche, le codage est une sorte de dépersonnalisation pour protéger l'identité des participants et des organismes impliqués

dans la recherche [1].

Renseignements personnels : Les renseignements personnels désignent toute information se rapportant à une personne physique et qui permettent de l'identifier, directement ou indirectement [3].

Renseignements d'identification directe : Renseignements permettant d'identifier une personne en particulier par des identificateurs directs (ex.: nom, adresse, numéro de téléphone, numéro d'assurance maladie, numéro de dossier CHUSJ, photo d'un patient). Certains renseignements en format texte (ex. : notes du médecin) ou des images (ex. : scan d'un patient) peuvent contenir des identifiants directs [1].

Renseignements d'identification indirecte : Renseignements qui peuvent vraisemblablement permettre d'identifier une personne par une combinaison d'identificateurs indirects (ex. : sexe, date de naissance, dates d'évènements (admission, diagnostic, procédure, congé), lieux (codes postaux, noms d'établissement de santé, lieu de résidence, caractéristique personnelle distinctive). Certains renseignements en format texte (ex. : notes du médecin) ou des images (ex. : scan d'un patient) peuvent contenir des identifiants indirects [1].

Renseignements anonymisés : Renseignements dont tous les identificateurs directs sont irrévocablement retirés et pour lesquels aucun code permettant une réidentification ultérieure n'est conservé. Le risque de réidentification de la personne à partir des identificateurs indirects restants est faible ou très faible [1].

3. Cadre légal

Les normes en matière de respect de la vie privée permettent l'utilisation et la divulgation restreintes des renseignements personnels aux fins des projets de recherche, pourvu que certaines exigences soient satisfaites.

Ce Cadre de gouvernance reconnaît l'importance du droit fondamental au respect de la vie privée et de la conformité aux normes applicables, notamment la Charte des droits et libertés de la personne, chapitre C-12 [4], le Code civil du Québec [5], la Loi sur les services de santé et les services sociaux, RLRQ chapitre S-4.2 (« LSSSS ») [6] et la Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels, *RLRQ chapitre A-2.1* (« LADOP ») [7].

La *Déclaration de principes des trois organismes sur la gestion des données numériques (2016*) [8] a également servi de référence à l'élaboration du présent document.

Ce cadre de gouvernance est cohérent avec les politiques et processus du CHUSJ, notamment:

- La Politique sur la sécurité de l'information du CHUSJ (POL-243) [9],
- La Politique sur la gestion des accès informationnels (POL-242-40) [10],
- La Politique sur les incidents de sécurité informationnelle (POL-242-50) [11],
- La Politique sur la confidentialité et l'accès au dossier (POL-2110),
- La Politique cadre sur la gestion des renseignements cliniques (POL-2101)

Finalement, l'architecture, la mise en œuvre et le maintien du fonctionnement de l'UnIC répondent aux directives ministérielles en matière de recherche et de sécurité informatique, incluant:

- La directive sur la cybersécurité [8,12],
- La directive sur l'utilisation sécuritaire des outils de collaboration par les médecins [13].
- Les termes et conditions d'utilisation des outils de collaboration [8,12,14].

4. Structure organisationnelle

La schématisation de la structure de fonctionnement de l'UnIC est présentée sous forme d'organigramme à l'annexe B

4.1 Mandats et composition des comités[1]

4.1.1 Comité de direction

Le Comité de direction a comme mandat principal de définir les orientations et les priorités de développement, notamment en termes d'intégration des systèmes d'information. Le Comité de direction approuve le Cadre de gouvernance du projet et veille à ce que les développements soient conformes aux besoins institutionnels dans le

[1] *Dans le présent document, les termes employés pour désigner des personnes sont pris au sens générique; ils ont à la fois valeur d'un féminin et d'un masculin.*

respect des échéanciers et du budget. Dans l'exercice de son mandat, le Comité de direction doit tenir compte des recommandations du Comité de gouvernance et résultats d'audit du responsable de la qualité et mettre en œuvre les priorités d'action. Le Comité de direction détermine le mandat et la composition du Comité de gouvernance de l'UnIC et du Comité des opérations.

Le comité est présidé par le Directeur de la Recherche (DR) et est composé des membres suivants: le Président-Directeur-Général Adjoint du CHUSJ, le Directeur des services professionnels (DSP), le Directeur de la Qualité, Évaluation, Performance et Éthique (DQEPE), le Directeur des Technologies, Ressources informationnelles et Génie Biomédicale (DTRIGBM), le représentant de la DTRIGBM responsable de l'extraction des données, le Président du Comité d'éthique à la recherche, le Directeur de la recherche clinique (DR), le Directeur de l'UnIC et le Coordonnateur de l'UnIC.

4.1.2 Comité de performance et gouvernance de l'UniC

Le Comité de performance et gouvernance a pour mandat d'établir le cadre de gouvernance et d'accompagner le projet tout au cours de son existence. Il doit, entre autres, offrir son soutien pour le développement des procédures entourant l'accès et définir les modalités d'appariement des données de l'UnIC avec celles des organisations ministérielles ou d'autres organisations.

Le Comité est composé des membres suivants: le Président du Comité d'éthique de la recherche (Président du Comité), le chef du Service des archives médicales, une représentante du Bureau des Affaires juridiques, un représentant de la Direction de la Qualité, Évaluation, Performance et Éthique (DQEPE), un représentant de la Direction des Technologies, Ressources informationnelles et Génie Biomédicale (DTRIGBM), l'officier responsable des actifs informationnels de la DQEPE, le Directeur de la recherche clinique (DR), un représentant de la pharmacie, un patient partenaire responsable du Bureau du Partenariat Patients-Familles-Soignants et représentant les patients et leur famille, le Directeur et le Coordonnateur de l'UnIC. D'autres membres peuvent être invités au besoin.

4.1.3 Comité des opérations

Le Comité des opérations de l'UnIC coordonne les travaux des équipes opérationnelles et veille à ce que les délais de développement et de mise en œuvre soient respectés tout en prenant en compte les obligations légales et réglementaires, les standards de qualité et de rigueur scientifique. Le Comité a également le mandat de s'assurer que les politiques, les procédures et les meilleures pratiques mises en place dans l'équipe de l'UnIC respectent le cadre de gouvernance et de gestion établi.

Le comité des opérations est présidé par le Directeur de l'UnIC et est composé des membres suivants: le Coordonnateur de l'UnIC, un ou plusieurs représentants de la DTRIGBM ainsi qu'un ou plusieurs représentants des chercheurs.

4.1.4 Directeur de l'UnIC

Le Directeur de l'UnIC assure la direction générale du projet, son suivi régulier, la prise des décisions, la gestion des risques, des finances et de l'échéancier. Ce responsable est nommé par le Président-Directeur-Général de l'hôpital. Le responsable dirige les équipes opérationnelles et veille à ce que celles-ci respectent les obligations légales et réglementaires, les standards de qualité et de rigueur scientifique. Il a également le mandat de s'assurer que les politiques, les procédures et les meilleures pratiques mises en place dans l'équipe de l'UnIC respectent le cadre de gestion établi.

Les tâches ou les responsabilités du Drecteur de l'UnIC pourraient être déléguées par le responsable à un membre du personnel si celui-ci possède l'expertise nécessaire pour mener à bien la tâche. Toute tâche ou responsabilité déléguée par le Directeur de l'UnIC est documentée.

4.1.5 Détenteur de l'information

Le détenteur de l'information est un employé désigné nommé par le Président-Directeur Général du CHU Sainte-Justine dont le rôle est de s'assurer de la sécurité de l'information et des ressources de l'UnIC. Cette personne a les compétences nécessaires à la gestion de la sécurité de l'information du projet. Il (ou elle) doit notamment planifier les activités de mise en place de la sécurité de l'information au sein du projet, veiller à l'application de la politique sur la sécurité de l'information et du cadre normatif de sécurité de l'information (CNSI) ainsi que veiller à la mise en œuvre de toute recommandation découlant d'une vérification ou d'un audit de sécurité [9].

4.1.6 Responsable de la qualité

Le Responsable de la qualité réalise les audits internes de conformité aux procédures et effectue le suivi auprès du Comité de performance et de gouvernance ainsi qu'au Comité de direction. Il s'agit d'un individu employé par le CHUSJ ne faisant pas partie de l'équipe de l'UnIC. Cette personne peut être le responsable de la qualité ou un autre employé du CHUSJ ayant les compétences requises pour effectuer des audits internes de conformité. Le Comité de direction peut également mandater un fournisseur à titre de responsable de la qualité.

4.1.7 Comité d’accès

Le Comité d’accès de l'UnIC veille au respect des accès aux données par les utilisateurs autorisés. Dans l’exercice de ses fonctions, ce comité peut être appelé à effectuer les tâches suivantes:

- Faire une évaluation préliminaire de la faisabilité des demandes d’accès avant l’obtention des approbations requises.
- Faire une évaluation des risques d’identification ou de réidentification des usagers du CHUSJ. Dans certains cas, l’analyse des sous-ensembles de données comportant un risque **élevé** de réidentification est réalisée par le scientifique de la donnée qui est attitré au projet.
- Déterminer les actions à effectuer sur les données pour réduire le risque d’identification des patients. Ces actions peuvent inclure entre autres, le masquage, la suppression, la pondération et l’agrégation des données.
- Déterminer les modalités d’accès aux données par les utilisateurs (ex. accès contrôlé dans l'environnement de l'UnIC). Le transfert et l’analyse des données de même que la diffusion des résultats d’analyse sont dépendants du risque de réidentification.
- Offrir un soutien scientifique aux chercheurs.

Ce comité est composé du Directeur et du Coordonnateur de l'UnIC, d’au moins un membre de l’équipe scientifique qui est attitré au projet et, au besoin, un administrateur du lac de données et*/*ou un membre de la Direction de la qualité, évaluation, performance et éthique (DQEPE). Toute information relative à une demande d’accès communiquée par le chercheur au personnel de l'UnIC est gardée confidentielle.

4.1.8 L’équipe de l’UnIC

L’équipe de l'UnIC est constituée d’experts en ingénierie des données massives, développement d'algorithmes et de solutions en recherche opérationnelle, gestion des plateformes technologiques et en science de données. Les membres de cette équipe sont des employés du CHUSJ ainsi que des collaborateurs externes. Tous les membres du personnel travaillent sous la supervision du Directeur de l'UnIC.

5. Données intégrées

Les systèmes à interfacer touchent à plusieurs services hospitaliers, incluant mais ne se limitant pas aux services des admissions, départs et transferts (ADT), de la pharmacie, des laboratoires, de l'imagerie médicale et de l'urgence. Les données à intégrer peuvent aussi provenir de banques de données recherche alimentées par les dossiers des usagers. D'autres systèmes à interfacer pourront s'ajouter en fonction des futurs besoins de la recherche et d'amélioration de la qualité, d'évaluation ou de suivi de la performance.

Les données brutes provenant de ces systèmes sont intégrées de façon prospective et continue dans leur format d'origine. Les données rétrospectives qui ont été collectées depuis la mise en fonction des systèmes sources seront aussi intégrées pour permettre aux chercheurs notamment, d'obtenir un portrait global de la trajectoire de soins de l'usager depuis la naissance (ou la première date d'admission), jusqu'à la vie adulte. Les données des différents systèmes sources peuvent être appariées en utilisant une clé d'indexation unique. L'ensemble des données intégrées à la plateforme de l'UI constitue des renseignements confidentiels au sens de la Loi sur les services de santé et les services sociaux.

6. Confidentialité et protection de la vie privée

- L'ensemble des renseignements hébergés dans l'environnement de l'UnIC, qu'ils soient sur une forme identifiable ou dépersonnalisée, constituent des renseignements personnels confidentiels au sens de la Loi sur les services de santé et les services sociaux et au sens de la Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels.

- La conservation des renseignements personnels est autorisée et nécessaire au bon fonctionnement du service offert par la plateforme de l'UI.

- Les chercheurs dans le domaine de la santé ont généralement besoin de données interreliées à l'échelle individuelle pour leurs analyses.

- Malgré le fait que les identifiants directs (nom, adresse, numéro d'assurance sociale, etc.) soient retirés, les renseignements qui demeurent pourraient permettre l'identification indirecte de la personne et doivent être gérés comme s'ils étaient des identifiants directs (renseignements personnels).

- Le droit à la vie privée des patients ainsi qu'à la confidentialité de leurs données personnelles est primordial pour ce projet.
- La divulgation non autorisée de ces données aurait potentiellement des répercussions réputationnelles, juridiques et financières à long terme pour le CHUSJ et potentiellement pour les personnes touchées par une telle fuite.

Tenant compte de ces énoncés, les principes de protection de la vie privée et de la confidentialité des données constituent l'assise de l'architecture de la plateforme et des mesures de sécurité s'y rattachant. Ces principes sont les suivants: 1) l'application du privilège minimal et de séparation des tâches, 2) la gestion des données en fonction du niveau de risque et du degré de sensibilité; 3) l'évaluation des risques par projet de recherche.

6.1 Privilège minimal et séparation des tâches

L'UnIC est constitué d'une équipe pluridisciplinaire dont les membres sont appelés à effectuer différentes tâches en lien avec la mise en place et le maintien de la plateforme de données. Dans le but de préserver la confidentialité des données des usagers du CHUSJ, des mesures sont mises en place pour limiter les privilèges d'accès aux stricts besoins. En premier lieu, seuls les membres de l'équipe ayant signé l'engagement à la confidentialité spécifique à l'UnIC et qui œuvrent au développement et au bon fonctionnement de la plateforme de l'UnIC ont un accès sécurisé aux données dans l'environnement de l'UnIC, conformément à la Politique sur la gestion des accès informationnels du CHUSJ (POL-242-40) [10]. Ce personnel est hautement qualifié en gestion de données, programmation et en conception, développement et évaluation de plateforme informatique de données massives.

Les individus mandatés ont soit un profil d'administrateur ou de scientifique de la donnée avec les droits d'accès qui correspondent à ces profils. En tant qu'utilisateurs, ces individus doivent utiliser les systèmes et accéder à l'information dans le respect des accès qui leur sont accordés. Une journalisation des utilisations et des accès à la plateforme de l'UnIC constitue un des moyens de veille et détection des atteintes à la confidentialité. Conformément aux règles du Réseau de la santé et des services sociaux, tout le personnel attitré au projet doit respecter les modalités d'accès au réseau du CHUSJ (incluant la gestion de mots de passe) et prendre toutes les mesures appropriées pour protéger les renseignements confidentiels contre tout vol, perte, interception, utilisation ou divulgation non autorisée, en ayant notamment recours à des mesures de

protection, à des méthodes et à des systèmes conformes aux standards des meilleures pratiques de l'industrie.

À la base, une compartimentalisation de la plateforme en trois zones en fonction des niveaux de risque permet l'application du principe de privilège minimal et séparation des tâches (Annexe A: Architecture de la plateforme), limitant ainsi les accès aux renseignements personnels à un nombre restreint d'individus. Les permissions pour l'accès aux données à caractère personnel ne sont octroyées qu'aux administrateurs. Ceux-ci sont responsables de la gestion des renseignements personnels et des procédures s'y rattachant, incluant le processus de codage et de dépersonnalisation.

Les scientifiques de la donnée sont responsables de toutes les fonctionnalités liées à l'extraction, l'organisation et à l'analyse des sous-ensembles de données dépersonnalisés. **Les jeux de données sont préparés pour un seul projet de recherche ou un seul projet d'amélioration de la qualité, d'évaluation ou de suivi de la performance.**

L'extraction des données des systèmes sources se fera sous la responsabilité de l'équipe de la Direction des Technologies, Ressources informationnelles et Génie biomédicale du CHUSJ (DTRIGBM). La sécurité des données dans l'espace temporaire de stockage est donc à la charge de la DTRIGBM. Les membres de l'équipe de l'UnIC n'ont et n'auront jamais accès aux systèmes sources du CHUSJ. Une fois les données extraites, celles-ci sont automatiquement importées de manière cryptée (processus de chiffrement) dans un espace sécurisé de l'infrastructure de l'UnIC (la zone rouge de l'Annexe A) accessible qu'à un nombre restreint d'administrateurs de l'équipe de l'UnIC. C'est dans cet environnement sécurisé que les données sont classées et que des programmes informatiques sont exécutés pour réconcilier les identifiants, créer un identifiant unique pour chaque patient propre à l'UnIC et exécuter la procédure de dépersonnalisation des données.

6.2 Gestion des données en fonction du niveau de risque

Le classement des renseignements au préalable est effectué [11,15] pour définir les balises d'accès et de gestion des différents types d'information contenus dans le lac de données et sert de guide pour l'évaluation du risque à la réidentification. Les renseignements sont classés en trois principaux types.

- Renseignements d'identification directe

- Renseignements d’identification indirecte
- Renseignements anonymisés

Dans la zone hautement sécurisée de la plateforme, les renseignements identificatoires les plus susceptibles d’identifier une personne seront identifiés, physiquement séparées des données de santé pour la recherche et soumises à des mesures techniques et organisationnelles afin de garantir la protection de la vie privée. Ces renseignements constituent le Registre-Patients et incluent notamment le prénom, nom, préfixe et suffixe de l’usager ou des personnes avec un lien de parenté, l’adresse, numéro de téléphone, courriel, numéro d’assurance-maladie, numéro de dossier, date de naissance et titre d’emploi.

Les données de santé seront modifiées pour en réduire le risque de réidentification. Ceci inclut la suppression des renseignements identificatoires et la création d’un code pour protéger l’identité des patients. Ce sont les données dépersonnalisées qui seront rendues accessibles aux scientifiques de la donnée. Brièvement, l'UnIC crée un identifiant unique pour chaque patient qui remplacera les renseignements d’identification directe et sera utilisé pour créer les liens entre les données des différents systèmes sources. Ces identifiants, de même que tous autres identifiants seront encryptés.

6.3 Évaluation des risques par projet

Les risques liés aux projets nécessitant les données hospitalières hébergées dans le lac de l'UnIC sont évalués en tenant compte du caractère identifiable des données, de la vulnérabilité des sujets en tant qu’individus ou collectivité et de la sensibilité des données [16]. La portée de l’utilisation des données est aussi prise en compte lors de l’évaluation des risques éventuels. Dans leur demande d’accès, les chercheurs doivent être précis quant à la portée des utilisations potentielles et aux mécanismes par lesquels des personnes à l'extérieur de l’équipe de recherche pourraient obtenir les données.

Une analyse de risque de réidentification est effectuée pour chaque projet par le Comité d’accès. Cette évaluation est réalisée dans le contexte spécifique du projet faisant l’objet d’une demande d’accès et permet d’attribuer à un jeu de données l’un ou l’autre des niveaux de risque suivant:

- **Risque élevé** : Le jeu de données comprend des éléments avec identifiants directs ou suffisamment d’identifiants indirects pouvant être utilisés pour identifier un individu;

- **Risque modéré**: Le jeu de données comprend majoritairement des données d'identification indirecte (dates de naissance, données géographiques, dates d'admission, congés, procédures spécifiques) pouvant être partiellement exposées à la réidentification.
- **Risque faible** : Le jeu de données contient des données ne contenant aucun élément avec identifiants directs et les éléments avec identifiants indirects ne sont pas suffisamment nombreux ou ont été manipulés pour assurer un niveau acceptable de risque à la ré- identification.

Dans la plupart des cas, des identifiants directs peuvent être facilement supprimés sans perturber la recherche, tant qu'il reste possible d'apparier les données provenant de différents systèmes sources, car l'identité des individus n'est pas importante pour la recherche. Cependant, la transmission d'éléments comportant un risque de ré-identification modéré ou élevé est possible dans les situations où le chercheur détient les autorisations requises. Dans tous les cas, seules les données autorisées seront transférées aux chercheurs.

Certaines données qui comportent un risque plus élevé de réidentification peuvent aussi nécessiter un traitement par les scientifiques de la donnée de manière à réduire le ce risque. Ceci peut être le cas de certaines données qui peuvent être masquées ou de données spécifiques (date de naissance, date d'évènement) qui peuvent être généralisées (groupe d'âge, année de l'évènement, âge à l'évènement). En fonction des questions posées dans les projets, les décisions concernant la technique la plus pertinente à utiliser et les variables à masquer font en partie appel au jugement du scientifique de la donnée attitré au projet.

Lors de l'utilisation des jeux de données avec des risques élevés ou modérés, il pourrait y avoir des restrictions/conditions, déterminées à l'avance par le Comité d'accès, sur la diffusabilité des données en cours ou suite à la recherche (non diffusable, sous condition, librement).

7. Sécurité de l'information

Le CHUSJ s'emploie à protéger les services et l'information de l'organisation.

Plusieurs lois, règlements, directives ou politiques encadrent et régissent l'utilisation et la gestion de l'information, notamment la politique sur la sécurité de l'information [9] qui met en place une gouverne claire de la sécurité de l'information.

Plus spécifiquement, les objectifs du CHU Sainte-Justine en matière de sécurité de l'information sont d'assurer :

- Le respect de la vie privée des individus, notamment, la confidentialité des renseignements personnels relatifs aux patients, aux participants à des études de recherche et aux intervenants du Réseau de la santé et des services sociaux.
- La disponibilité, l'intégrité et la confidentialité de l'information à l'égard de l'utilisation qui en est faite au CHU Sainte-Justine.
- Le respect des mesures de sécurité concernant l'utilisation des actifs informationnels.
- La conformité aux lois et règlements applicables ainsi qu'aux directives, normes et orientations gouvernementales [2].

Les mesures physiques mises en place dans le contexte de l'UnIC visent à assurer la confidentialité à tous les niveaux du cycle de vie des données qui comprend l'extraction, l'intégration, l'utilisation, la diffusion, la conservation et l'élimination des données. Celles-ci adhèrent aux politiques et processus du cadre normatif de sécurité de l'information du CHUSJ et répondent aux directives ministérielles en matière de sécurité informatique.

Un processus de catégorisation des risques permettant d'énumérer et d'évaluer la gravité des impacts que pourrait provoquer un bris de sécurité a été effectué au cours de la planification des mesures et objectifs de sécurité pour le projet de l'UnIC. L'analyse, qui s'est penchée sur la disponibilité, l'intégrité et la confidentialité des données, a permis de mettre en évidence les barrières de sécurité et la gravité des incidents de sécurité. Compte tenu de la gravité d'un incident mettant en cause la confidentialité des données du système, la mise en œuvre de mesures particulières au niveau de l'accès et de la manipulation des informations doit permettre de satisfaire à un besoin d'intégrité et confidentialité.

L'élaboration et l'organisation de l'architecture générale du système de sécurité reposent sur le concept de défense en profondeur selon lequel la sécurité ne doit pas reposer sur une seule technologie ou un seul produit de sécurité mais plutôt sur un ensemble cohérent de stratégies qui doivent être surveillées, protégées et qui doivent bénéficier d'un plan de réaction et mitigation en cas d'incident.

7.1 Mesures matérielles et technologiques

Des mesures matérielles et technologiques visant à assurer la sécurité de l'information seront mises en place conformément aux politiques en vigueur au CHUSJ.

Le Lac de données de l'UnIC et tous les entrepôts de données et sous-ensembles en découlant sont entreposés sur des serveurs dédiés et protégés au sein des infrastructures du CHUSJ. Le local où sont hébergés les serveurs est verrouillé avec une stricte gestion d'accès à l'environnement. Ainsi, les actifs informationnels de l'UnIC sont soumis aux mêmes standards de sécurité que les données détenues dans les dossiers des usagers du CHUSJ. Seules les personnes responsables de l'organisation et de la maintenance des serveurs dédiés à l'UnIC moyennant une authentification par badge peuvent accéder à l'espace sécurisé des serveurs.

L'architecture technologique exploite plusieurs techniques de sécurité afin de réduire l'exposition du système aux différentes menaces, notamment:

- Un découpage des zones en fonction des risques, du profil des utilisateurs et des étapes du cycle de vie des données (voir schéma de l'architecture, Annexe A). Dans ce cadre, sont rattachées des règles d'accès et de passage de l'une à l'autre des zones;
- Une journalisation des accès des utilisations de la plateforme;
- Une gestion des accès via un gestionnaire d'accès "open source";
- L'application de l'authentification forte (double facteurs) lorsqu'un utilisateur veut se connecter à partir d'un point d'accès situé à l'extérieur du RSSS, conforme aux standards du MSSS;
- L'encryption des données en transit et au repos;
- Une sauvegarde quotidienne des données;
- Une vérification des sauvegardes semestriellement;
- Une documentation de la gestion de mots de passe système;
- Une mise à jour récurrente des systèmes d'exploitation.

7.1.1. Évaluation de la vulnérabilité et tests d'intrusion

Les vérifications représentent une composante importante du programme global de sécurité de l'information de l'UnIC. Elles visent à s'assurer du respect des pratiques exemplaires et à mesurer la conformité avec l'ensemble des politiques, procédures et pratiques de sécurité de l'information du CHUSJ. Ces vérifications servent entre autres à évaluer la capacité de l'UnIC à protéger l'information et les systèmes de traitement de l'information contre les menaces et vulnérabilités. Un audit externe peut être réalisé sur une base annuelle et des audits locaux peuvent être réalisés sur une base plus fréquente.

Les évaluations de la vulnérabilité et les tests d'intrusion de l'infrastructure et de certaines applications constituent une composante importante du programme de vérification de l'UnIC. La gestion des vulnérabilités est un exercice mensuel en raison du dynamisme de l'environnement des menaces. Des exercices d'intrusion sont menés sur une base annuelle et au niveau local sur une base plus fréquente. Toutes les recommandations formulées dans le cadre des vérifications font l'objet d'un suivi et les mesures appropriées sont prises le cas échéant.

7.2 Mesures organisationnelles

7.2.1 Gestion des permissions

Les membres du personnel de l'UnIC, qu'ils aient un profil d'administrateur ou de scientifique de la donnée, doivent utiliser les systèmes et accéder à l'information dans le respect des accès qui leur sont accordés. Chaque utilisateur a un identifiant unique et le mécanisme d'authentification permet de vérifier l'identité déclarée de cet utilisateur. De plus, les droits d'accès de ces individus sont revus trimestriellement. Conformément aux règles du Réseau de la santé et des services sociaux, tout le personnel attitré au projet doit respecter les modalités d'accès au réseau du CHUSJ (incluant la gestion de mots de passe) et prendre toutes les mesures appropriées pour protéger les renseignements confidentiels contre tout vol, perte, interception, utilisation ou divulgation non autorisée, en ayant notamment recours à des mesures de protection, à des méthodes et à des systèmes conformes aux standards des meilleures pratiques.

7.2.2 Formation exigée

Tous les membres du personnel de l'UnIC doivent suivre la formation en éthique de la recherche clinique offerte par le CHUSJ. Cette formation doit être renouvelée tous les 3 ans. Ce cours se penche sur les grands principes régissant la recherche impliquant des humains. Un certificat attestant la complétion de cette formation est exigé avant d'accorder un accès aux ressources de l'UnIC. Par ailleurs, tous les membres doivent aussi être sensibilisés régulièrement à l'importance de maintenir la confidentialité des renseignements personnels sur la santé et doivent se familiariser et se tenir à jour avec certaines politiques et procédures du cadre normatif de sécurité du CHUSJ.

7.2.3 Engagements des membres de l'équipe de gestion des données

Une enquête de sécurité sur le personnel est requise pour les membres du personnel de l'UnIC. De plus, tous les membres du personnel doivent signer un engagement à la confidentialité détaillé et spécifique au projet stipulant, entre autre, qu'ils s'engagent à prendre toutes les mesures requises pour protéger la confidentialité des renseignements personnels et à ne les dévoiler à quiconque, que ce soit sous une forme verbale ou écrite, sauf à un autre membre de l'équipe du projet UnIC assujetti à un engagement à la confidentialité ou à un utilisateur secondaire ayant reçu les autorisations nécessaires. Les membres comprennent par ailleurs que le CHUSJ exerce une surveillance continue des accès et utilisations des données et des outils logiciels dans l'environnement de l'UnIC et que tout manquement à leur engagement peut mener à des sanctions.

8. Mise à disposition des données

8.1 Conditions d'accès

8.1.1 Dans le cadre d'un projet de recherche

Pour qu'elle soit autorisée, l'utilisation des données hébergées dans la plateforme de l'UnIC doit être conforme au Cadre réglementaire pour la recherche avec des participants humains adopté au CHUSJ.

Les projets de recherche pour lesquels toutes les autorisations requises ont été obtenues peuvent faire l'objet d'une demande d'accès aux données de l'UnIC. Ces autorisations comprennent notamment :

- L'approbation du comité d'éthique de la recherche,
- Le consentement des participants/de leur représentant légal ou l'autorisation du Directeur des services professionnels (en absence du consentement des participants/de leur représentant légal),
- L'approbation du comité de convenance de la Direction de la recherche,
- L'autorisation de la personne formellement mandatée pour autoriser la recherche au CHUSJ,
- Autres autorisations d'organismes détenteurs de données, si applicables.

L'instigateur de la demande d'accès doit avoir un statut de chercheur au CHUSJ pour pouvoir faire une demande d'accès aux données de l'UnIC. Pour les projets provenant du

milieu académique externe au CHUSJ ou pour les projets provenant de l'industrie, une collaboration doit être établie avec un chercheur du CHUSJ préalablement à la demande d'accès aux données.

Les documents requis pour effectuer une demande d'accès aux données sont:

- Le protocole de recherche du chercheur, incluant au besoin, un plan de gestion de données,
- Le CV du chercheur,
- La preuve des différentes autorisations requises.

8.1.2 **Dans le cadre d'un mandat d'amélioration de la qualité, d'évaluation ou de suivi de la performance**

Afin de mener à terme des projets d'amélioration des soins et services ou de suivre la performance de l'établissement, les gestionnaires, employés et médecins de l'hôpital, sous la supervision de l'équipe Performance, relevant de la Direction qualité, évaluation, performance et éthique (DQEPE), pourront utiliser les données de l'UnIC. Pour être recevable par l'UnIC, une demande d'accès relative à un projet d'amélioration de la qualité, d'évaluation ou de suivi de la performance doit préalablement être déposée à la DQEPE pour fin d'évaluation de la pertinence via le formulaire web de demande de service.

La nécessité d'utiliser les ressources de l'UniC pour répondre aux besoins des projets d'amélioration de la qualité, évaluation ou suivi de la performance sera évaluée par la DQEPE. La DQEPE sera, par ailleurs, responsable de la gestion de la procédure d'accès aux données en amont, notamment de s'assurer que toute demande ait obtenu les autorisations requises, incluant notamment l'autorisation du Directeur des services professionnels (DSP) et du Conseil des médecins, dentistes et pharmaciens (CMDP), le cas échéant.

La nature des mandats traités par l'équipe Performance, DQEPE, exige fréquemment des positionnements rapides à l'intérieur de délais restreints. Conséquemment, certaines mesures pourraient être mise en place pour permettre un accès rapide aux sources de données pour soutenir la prise de décision organisationnelle. Par exemple, des ensembles de données pourraient être rendus disponibles au personnel préautorisé dans l'espace chercheur de l'UnIC et rafraîchis régulièrement pour accélérer l'accès aux données pour des fins de reddition de comptes ou demandes récurrentes.

8.2 Traitement d’une demande d’accès

Les demandes d’accès aux données de l'UnIC sont traitées par le comité d’accès suite à l’obtention de toutes les autorisations requises.

8.2.1 Délai d’accès

Une confirmation de la soumission d’une demande d’accès sera envoyée dans les deux jours ouvrables suivant la réception de la demande. Plusieurs facteurs peuvent influencer les délais d'accès, notamment la disponibilité de l’instigateur de la demande pour répondre aux questions de l’analyste, la complexité de la demande (ex.: interrogation complexe de la base de données pour évaluer la faisabilité) et l’obtention d’autres autorisations requises. La demande sera toutefois traitée de façon diligente.

8.2.2 Coût des services

Les demandes d’accès aux données et de services à l'UnIC sont sujettes à un recouvrement de coût selon le modèle des plateformes du CHUSJ. L’évaluation de la faisabilité détaillée permet de quantifier les coûts qui seront spécifiés dans l’Entente d’accès aux données. Les estimés sont sujets à changement advenant que des modifications soient apportées au projet.

L’accès aux données de l’UnIC requis pour le traitement de demandes dans le cadre d’un mandat d’amélioration de la qualité, d’évaluation ou de suivi de la performance soutenues par l’équipe Performance ne feront pas l’objet de facturation.

8.3 Utilisation des données

8.3.1 Entente d’accès

Suite à l’obtention de toutes les autorisations requises, l’Entente d’accès aux données doit être dûment signée par l’utilisateur, le Directeur de l'UnIC et le Directeur de la recherche du CHUSJ, pour permettre le transfert et l’utilisation des données. Il est à noter que la portée de cette entente d’accès se limite aux projets de recherche et n’inclut pas les projets portant sur l’amélioration de la qualité, l'évaluation et le suivi de la performance.

Cette entente réitère les responsabilités de l'utilisateur vis-à-vis la protection de la vie privée des usagers du CHUSJ et la confidentialité des données, notamment que celui-ci s'engage à garder confidentiel tout renseignement personnel d'un participant dont il pourrait prendre connaissance, même involontairement, de porter à la connaissance du comité d'accès de l'UnIC tout changement significatif apporté au protocole du projet de recherche et toute situation susceptible d'affecter la sécurité ou la confidentialité des données.

L'entente peut aussi prévoir, dans certains cas, que le chercheur doit valider son interprétation des données avec la personne pilote du système source utilisé.

8.3.2 Découverte d'une erreur fortuite

Des erreurs peuvent exister dans les ensembles de données fournis par l'UnIC. Ces erreurs peuvent être dues à une mauvaise manipulation des données, telles qu'un appariement erroné entre deux sources de données, la présence injustifiée de renseignements personnels ou la présence de doublons. L'utilisateur est responsable d'informer l'équipe de l'UnIC advenant la découverte de telles erreurs fortuites dans le sous-ensemble de données afin que les erreurs soient documentées et que les corrections nécessaires soient apportées.

Par ailleurs, si l'utilisateur détecte une incohérence dans les données pouvant avoir un impact sur la prise en charge clinique d'un usager, par exemple, une erreur possible dans le diagnostic d'un patient, celui-ci devra aviser l'UnIC pour qu'une évaluation du cas soit effectuée. Dans le cas où l'erreur est confirmée, un signalement doit être fait auprès du Directeur des services professionnels (DSP).

8.3.3 Propriété intellectuelle

Dans le cas où une collaboration de recherche est établie autour du projet utilisant les données de l'UnIC, l'utilisateur s'engage à respecter les règles applicables en matière de propriété intellectuelle selon la collaboration établie.

Le CHUSJ ne peut revendiquer de droits de propriété intellectuelle du seul fait d'avoir rendu accessible des données de l'UnIC ayant contribué à la réalisation de découvertes, inventions ou œuvres.

En aucun cas, l'utilisateur n'a de droits de propriété intellectuelle sur les données extraites de la base de données, ni le droit de commercialiser les données extraites de l'UnIC.

8.3.4 Publications de résultats

Les utilisateurs des données de l'UnIC sont invités à publier les résultats découlant de leur recherche afin que la communauté scientifique et la population en générale puissent en bénéficier. Toute publication ou présentation utilisant les données de l'UnIC doit:

- Mentionner que les données utilisées dans l'étude/présentation proviennent de l'UnIC du Centre hospitalier universitaire Sainte-Justine (CHUSJ);
- Documenter et décrire les sources de données utilisées;
- Inclure une clause de reconnaissance pour les fournisseurs des systèmes sources ou tout autre intervenant ayant permis la collecte ou l'interprétation des données. Ces clauses, fournies par l'UnIC , seront ajustées en fonction des sources de données utilisées;
- Indiquer que toutes les mesures pour préserver la confidentialité des données des usagers du CHUSJ et de leur famille ont été mises en œuvre dans le contexte de l'étude.

L'utilisateur est responsable de remettre au comité d'accès l'UnIC une copie du manuscrit rapportant les résultats du projet dans un délais de 30 jours ouvrables avant la date prévue de publication, afin d'en évaluer les risques d'identification et de stigmatisation et de s'assurer que la publication est cohérente avec le projet approuvé. De plus, l'utilisateur doit transmettre un résumé du projet de recherche et des résultats, en langage accessible au grand public, pour fins de communication sur le site internet du CHUSJ. Cette obligation se limite aux publications dans les revues scientifiques ou lors de congrès scientifiques, mais ne s'applique pas aux rapports internes.

L'utilisateur s'engage à reconnaitre la contribution d'un membre de l'équipe de l'UnIC lorsque celle-ci respecte les règles en matière de reconnaissance d'auteurs sur les publications [17].

L'UnIC se réserve le droit de contacter l'utilisateur, après la soumission du rapport afin de collecter des informations additionnelles visant à mesurer les retombées du projet (publications, présentations à des conférences, etc.).

8.3.5 Dépôt des données dans un registre public

Lors de la publication, il est possible que certaines revues scientifiques exigent que les données utilisées pour générer les résultats de recherche soient mises à disposition dans

un registre public. Dans de telles circonstances, le jeu de données ne doit pas être déposé dans un registre public de données sans l'autorisation formelle du comité d'éthique de la recherche.

8.3.6 Retour de données

Il est possible que l'utilisation des données génère des nouvelles données qui puissent être d'intérêt pour d'autres utilisateurs (ex: calcul d'un score de risque). Dans ces cas-ci, les utilisateurs seront invités à soumettre et rendre accessible leurs ensembles de données dérivées. Les exigences concernant le retour de telles données seront documentées dans le contrat d'accès. Par exemple, le contenu et le format des données ainsi que les critères entourant la documentation du code, des logiciels ou des analyses ayant permis de générer les nouvelles données devront être décrits dans l'entente d'accès.

8.3.7 Fin de l'accès et destruction des données

À l'arrivée du terme de la période pour laquelle l'accès a été approuvé, l'utilisateur doit en informer le comité d'accès de l'UnIC et respecter les exigences relatives à ce qui doit être fait des données concernées. Aucune copie des données ne doit être conservée par l'utilisateur. Au terme de l'entente, l'utilisateur devra, par conséquent, attester par écrit qu'il n'a pas conservé de copie des données en fournissant une attestation de destruction par écrit. Il est à noter que le code qui a généré les jeux de données et les copies de ces données seront conservés par l'UnIC pour utilisation future.

En ce qui concerne les projets de recherche, les ensembles de données sur lesquels ont travaillé les chercheurs seront conservés dans un répertoire propre au projet pour une durée minimale de 7 ans ou aussi longtemps que prévu par les exigences légales et contractuelles des organismes subventionnaires et partenaires, conformément aux exigences de l'Université de Montréal en matière de conservation de données de recherche [18].

9. Contrôle de la qualité et de la conformité

Malgré la mise en place d'un lac de données permettant la mise en commun et l'exploitation des données hospitalières, il n'en demeure pas moins que la qualité des recherches découlant de l'exploitation de l'UnIC est tributaire de la qualité des données collectées par les systèmes sources. Or, l'utilisation de ces données pour des fins de recherche ou d'amélioration de la qualité comporte certaines contraintes, notamment:

- Des contraintes techniques liées au format hétérogène et non structuré des données;
- Des contraintes d'interopérabilité liées à une adhésion variable des systèmes d'information hospitaliers à des terminologies comprenant des codes et des libellés;
- Des contraintes de qualité liées aux erreurs de codages ou autres.

Dans une perspective d'amélioration continue, l'équipe de l'UnIC mettra en place des procédures et mécanismes pour assurer une gestion des données qui adhère aux principes FAIR[2]. Ceci inclut, notamment:

- La mise en place d'une procédure permettant d'évaluer l'intégrité des données lors de leur chargement.
- L'utilisation d'opérations visant à réparer, nettoyer, transformer et standardiser les données, notamment en les formatant convenablement, en repérant les erreurs de transposition et les doublons. Le code informatique utilisé pour effectuer ces opérations sera conservé et bien documenté par souci de transparence. Les procédures de transformation ou nettoyage des données pourront, au besoin, être validées par les pilotes de systèmes.
- La création des dictionnaires de données pour permettre de documenter le contenu des tables de données ainsi que la transformation des données. Le contenu de ces dictionnaires de données pourra, au besoin, être validé par les pilotes de systèmes.
- Un plan de classement de l'espace de stockage pour retrouver facilement les différentes données.
- Une convention de nommage pour identifier les différents types de données.
- Un fichier de métadonnées, basé selon le standard de métadonnées approprié, associé à chaque jeu de données pour les décrire et aider à leur recherche et à leur identification.

[2] https://www.go-fair.org

- L’ajout d'un identifiant unique et pérenne pour chaque jeu de données.

La traçabilité du cycle de vie d’une donnée sera ainsi assurée de son état initial, jusqu’à son état final, en passant par toutes les transformations qui ont eu lieu entre ces étapes.

L'UnIC veille à mettre en place un processus de traçabilité permettant de savoir exactement où sont stockées et traitées les données et quels sont les accès qui ont été accordés pour certains jeux de données.

L’ensemble de ces mécanismes et procédures feront l’objet d’audit interne de conformité.

10. Appariement avec sources externes

Le couplage des bases de données provenant d'autres organisations que le CHUSJ, crée de nouvelles possibilités pour la recherche, mais aussi de nouveaux risques d’atteinte à la vie privée. Lorsqu’approprié et seulement pour l’usage défini spécifié dans le projet approuvé, il est possible que les données de l'UnIC soient appariées ou mises en commun avec d’autres données. Toute liaison ou combinaison des données de l'UnIC avec d’autres sources de données doit être spécifiée dans la demande d’accès et les autorisations requises doivent être obtenues avant l’obtention de l’accès aux données.

Certaines banques de données constituées à des fins de recherche peuvent aussi être appariées aux données clinico-administratives du CHUSJ entreposées dans l'UnIC si le consentement du participant/de son représentant légal le permet et si toutes les autorisations requises ont été obtenues. Il est à noter que les banques de données constituées à des fins de recherche ont des règles d’accès qui leur sont propres qui doivent aussi être respectées.

11. Adoption, implantation et révision

Le Comité de direction de l'UnIC est responsable de l’adoption du cadre défini dans le présent document. Le Directeur de l'UnIC, avec l’aide du comité des opérations, a la responsabilité d’assurer l’implantation et l’application du cadre.

Le document sera revu par le Comité de gouvernance tous les deux ans ou selon tout événement qui modifierait le cadre légal ou réglementaire sur lequel il s’appuie pour s’assurer de sa cohérence avec la législation actuelle et les meilleures pratiques.

Historique des Révisions

Version	Entrée en vigueur	Résumé des modifications
1	2022-05-26	Version originale

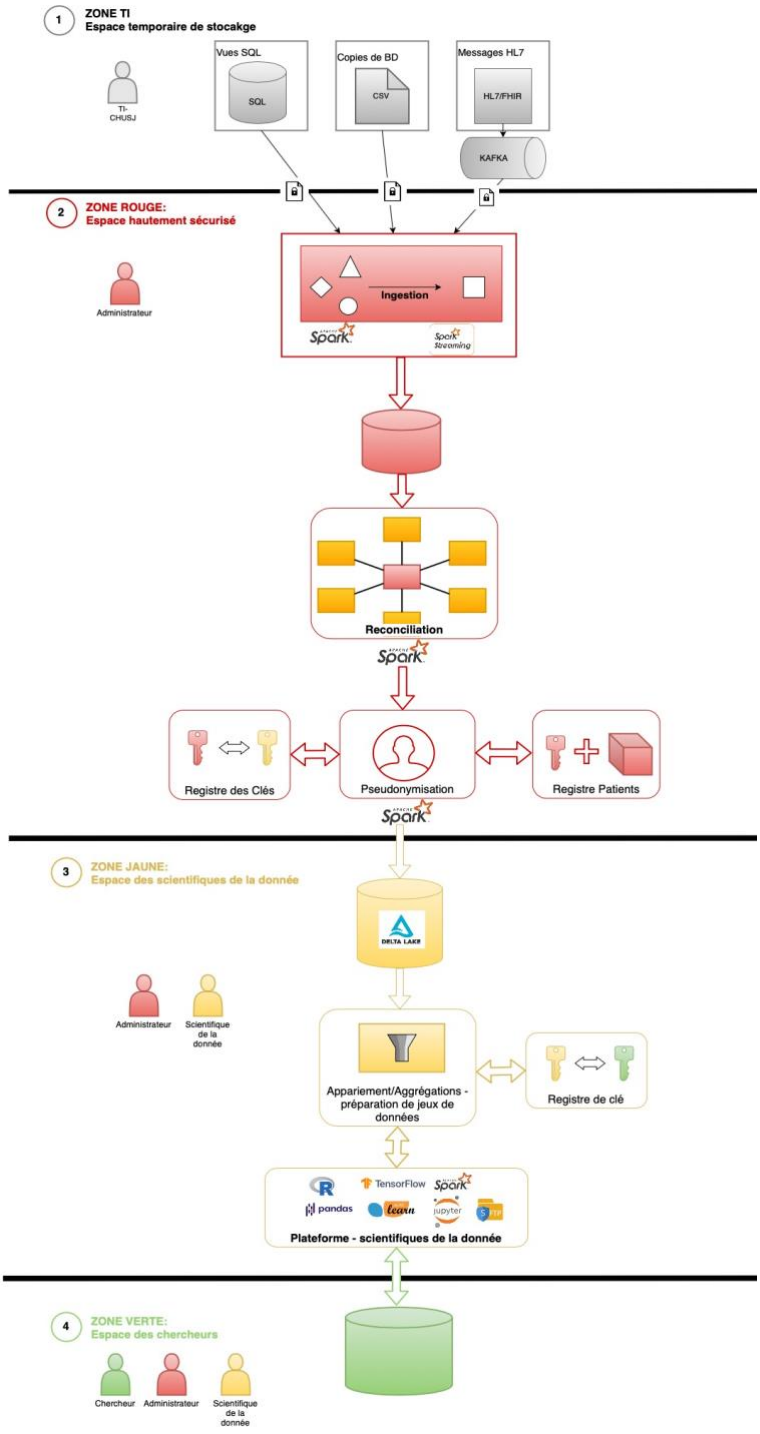
13. Rérérences

1. Portage Network Sensitive Data Expert Group on behalf of the Canadian Association of Research Libraries (CARL). *Sensitive Data Toolkit for Researchers Part 1: Glossary of Terms for Sensitive Data used for Research Purposes*. (2020).
2. Ministère de la Santé et des Services sociaux Direction générale des technologies de l'information. *MSSS-CDG01 Cadre de gestion de la sécurité de l'information*. (2015).
3. *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels*, chapitre A-2.1.
4. *Charte des droits et libertés de la personne du Québec,* chapitre C-12.
5. *Code civil du Québec*, L.Q. 1991, c.64.
6. *Loi sur les services de santé et les services sociaux*, chapitre *S-4.2*.
7. *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels*, chapitre *A-2. 1*.
8. Innovation, Sciences et Développement économique Canada, du sous-ministre du Canada & Direction générale des communications et du marketing. Déclaration de principes des trois organismes sur la gestion des données numériques. https://science.gc.ca/eic/site/063.nsf/fra/h_83F7624E.html.
9. Direction Qualité Performance. *Politique sur la sécurité de l'information*. (2018).
10. Direction Qualité Performance. *Politique sur la gestion des accès informationnels*. (2017).
11. Direction de la performance et responsable de la sécurité de l'information. *Politique sur les incidents de sécurité informationnelle*. (2019).
12. Ministère de la santé et des Services sociaux. *Directive sur la cybersécurité (MSSS-DIR03)*. (2018).

13. Ministère de la santé et des services sociaux. *MSSS-DIR04: Directive sur l'utilisation sécuritaire des outils de collaboration par les médecins*. (2020).
14. Ministère de la santé et des Services sociaux. *Termes et conditions d'utilisation des outils de collaboration*. (2020).
15. Committee on Strategies for Responsible Sharing of Clinical Trial Data, Board on Health Sciences Policy & Institute of Medicine. *Concepts and Methods for De-identifying Clinical Trial Data*. (National Academies Press (US), 2015).
16. Groupe d'experts sur les données sensibles (GEDS) du réseau Portage au nom de l'Association des bibliothèques de recherche du Canada (ABRC). *Boîte à outils pour les données sensibles --- destiné aux chercheurs. Partie 2: Matrice de risque lié aux données de recherche avec des êtres humains*. (2020).
17. International Committee of Medical Journal Editors. *Defining the Role of Authors and Contributors*. (2019).
18. 03000 - Recherche - Gestion de documents et des archives - Université de Montréal. *Archives Université de Montréal* https://archives.umontreal.ca/gestion-de-documents/regles-de-gestion/03000-recherche/#r03520.

14. Annexes

Annexe A: Schéma de l’architecture de l'UnIC



Trois couches de sécurité et types d'accès

Rouge

- ➔ Accès restreint aux administrateurs de l'UI
- ➔ Réconciliation, codage et dépersonnalisation des données

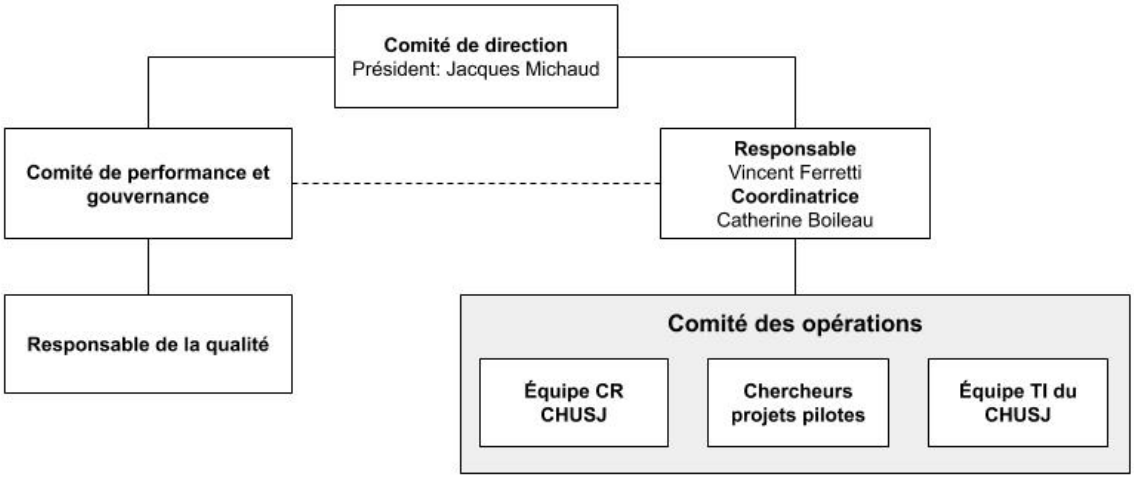
Jaune

- ➔ Accès restreint aux scientifiques de la donnée et aux administrateurs de l'UI
- ➔ Curation des données dépersonnalisées (codées) du lac,
- ➔ Préparation des jeux de données

Vert

- ➔ Espace sécurisé de travail pour chercheurs
- ➔ Accès par les chercheurs aux jeux de données autorisés dans espace sécurisé
- ➔ Identifiants uniques pour chaque jeux de données

Annexe B: Structure organisationnelle de l'UnIC



Annexe C: Documents reliés au Cadre de gouvernance

Titre du document	**Description**	**Status**
DOCUMENTS JURIDIQUES		
Engagement à la confidentialité	Engagement renforcé pour les pour les membres du personnel de l'UnIC	V 1.0
Entente d'accès aux données	Pour les projets de recherche	En développement
PROCÉDURES		
Procédure de dépersonnalisation		V 1.0
Procédure d'accès	Accès aux données pour les projets de recherche et les projets d'amélioration de la qualité	En développement
Procédure de fin de projet	Procédure administrative de fin de projet	En développement
Procédure pour la découverte d'une erreur fortuite		En développement
Analyse du risque de réidentification	Procédure d'évaluation du potentiel d'identifier un patient dans un jeux de données	En développement
Procédure d'ingestion de données	Procédure décrivant comment les données sont sélectionnées à partir de la base de données d'intégration et déposée dans la zone rouge du Lac de données	En développement
Procédure de cessation d'emploi	Procédure administrative couvrant les étapes suivant l'annonce de cessation d'emploi d'un membre de l'équipe	En développement
Document technique: Architecture du Lac de données		En développement