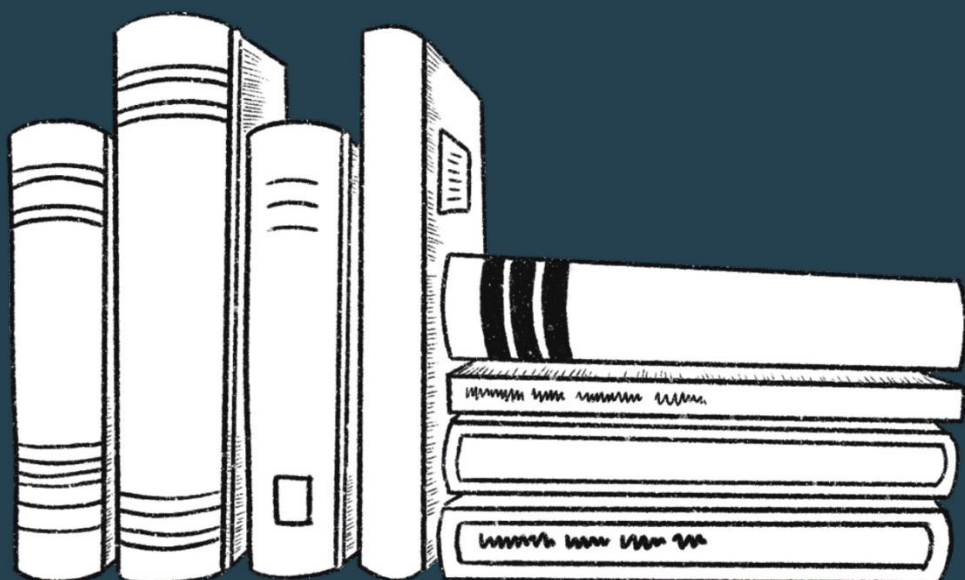


Conseil ontarien  
de la qualité de  
l'enseignement supérieur



**Plan directeur du PRC : Comment nous avons  
construit une infrastructure communautaire de  
données**

Jackie Pichette, Lorraine Valmadrid  
et Sally Landon

**Publié par le :**

## **Conseil ontarien de la qualité de l'enseignement supérieur**

**88, Queens Quay Ouest, bureau 2500  
Toronto (Ontario)  
M5J 0B8**

**Téléphone : 416 212-3893**

**Télécopieur : 416 212-3899**

**Site Web : [www.heqco.ca](http://www.heqco.ca)**

**Courriel : [info@heqco.ca](mailto:info@heqco.ca)**

### **Citer ce document comme suit :**

Pichette, J., Valmadrid, L. et Landon, S. (2022) *Plan directeur du PRC : Comment nous avons construit une infrastructure communautaire de données*. Toronto : Conseil ontarien de la qualité de l'enseignement supérieur



Conseil ontarien  
de la qualité de  
l'enseignement supérieur

Les opinions exprimées dans le présent document de recherche sont celles de l'auteur et ne reflètent pas nécessairement les vues ni les politiques officielles du Conseil ontarien de la qualité de l'enseignement supérieur ou d'autres organismes ou organisations ayant offert leur soutien, financier ou autre, dans le cadre de ce projet. © Imprimeur de la Reine pour l'Ontario, 2022.

# Table des matières

Liste des figures .....	3
Introduction .....	4
Structure de gouvernance.....	5
Autorité légale .....	7
Principes directeurs.....	9
Assemblage et stockage des données .....	13
Limites.....	15
Enseignements tirés et prochaines étapes.....	17
Références.....	19
Annexes .....	21
Annexe A : Autorité légale .....	22
Annexe B : Principes directeurs.....	24

## Liste des figures

Figure 1 Structure de gouvernance du PRC.....	6
Figure 2 Principes directeurs du PRC.....	12
Figure 3 Étapes de l'assemblage des données du PRC .....	15

## Introduction

Pour être efficaces, l'élaboration des politiques et la planification dans le secteur de l'éducation exigent des informations détaillées sur les expériences et les résultats des élèves, à la fois dans le temps et entre les établissements. Hélas, il n'est pas facile d'obtenir ce type d'information en Ontario (Gallagher-Mackay, 2017; Robson, 2021). C'est d'ailleurs cette lacune quant à l'information disponible sur le parcours scolaire des élèves qui a motivé la création du Partenariat de recherche communautaire (PRC) de Hamilton, une coalition de six organismes dont l'engagement consiste à favoriser la réussite de la population étudiante à Hamilton grâce à un réseau sécurisé d'échange de données.

Le présent document décrit le processus et les conditions qui ont régi la création et le déploiement d'un projet pilote du PRC. Il propose un plan directeur à d'autres communautés et établissements d'enseignement qui souhaitent s'inspirer des membres du PRC et créer des ensembles de données similaires à des fins de recherche et d'amélioration. Tout au long du présent document, nous détaillons l'évaluation de l'impact sur la protection de la vie privée qui a guidé notre approche et les étapes de l'assemblage de l'ensemble de données, y compris un processus de dépersonnalisation en deux étapes, le transfert et la fusion des données. Nous discutons ensuite des limites inhérentes à notre processus ainsi que de nos plans pour résoudre ces difficultés dans les années à venir. Le premier ensemble de conclusions tirées du PRC est présenté dans notre rapport complémentaire, [\*Le pouvoir des données connectées : Tracer le parcours des élèves qui entreprennent et terminent des études postsecondaires à Hamilton.\*](#)

Le PRC de Hamilton regroupe deux conseils scolaires locaux : le Hamilton-Wentworth District School Board (HWDSB) et le Hamilton-Wentworth Catholic District School Board (HWCDSD); deux établissements d'enseignement postsecondaire : l'Université McMaster (McMaster) et le Collège Mohawk d'arts appliqués et de technologie (Mohawk); et deux organismes publics axés sur la résolution des problèmes systémiques en éducation : la Hamilton Community Foundation (HCF) et le Conseil ontarien de la qualité de l'enseignement supérieur (COQES). Les partenaires du secteur de l'éducation (HWDSB, HWCDSD, McMaster et Mohawk) fournissent des données et leur expertise en tant que prestataires de services et de programmes d'éducation. Le COQES et la HCF agissent en tant que cofacilitateurs, en proposant une orientation stratégique et un soutien continu à la gestion du projet. La HCF est le principal point de contact du partenariat, et travaille en collaboration directe avec les partenaires pour résoudre les problèmes et maintenir le projet sur la bonne voie. Le COQES collabore avec la HCF et les partenaires pour réaliser des recherches, mobiliser les connaissances et conseiller les décideurs provinciaux, et il assure également le financement du PRC.

Les membres du PRC partageaient trois objectifs. Le premier consistait à mieux comprendre le parcours éducatif des élèves de Hamilton. Les données sont souvent limitées à chaque établissement, de sorte que l'on sait peu de choses sur les élèves après leur départ d'un établissement (par exemple, d'une école secondaire) et avant leur entrée dans un autre (par exemple, dans un collège). Cela rend presque impossible l'identification des facteurs systémiques qui affectent les trajectoires d'apprentissage et contribuent aux inégalités entre les élèves. Le deuxième objectif était de combler cette lacune dans les informations disponibles en élaborant et en testant un mécanisme de partage des données : nous avons utilisé le numéro d'immatriculation scolaire de l'Ontario (NISO) pour relier les données administratives au niveau individuel entre plusieurs conseils scolaires, partenaires du postsecondaire et centres de demandes d'admission. Chaque élève se voit attribuer un NISO unique qui le suit tout au long de son parcours éducatif, de la maternelle au secondaire et, s'il le souhaite, jusqu'au niveau postsecondaire. En utilisant le NISO comme clé, les partenaires du PRC ont cherché à relier les données administratives des établissements pour produire un ensemble de données à des fins de recherche. Notre objectif final était d'instaurer la confiance entre les membres du PRC et d'établir des procédures pour assurer une recherche continue et collaborative dans notre communauté – tout en inspirant d'autres communautés à poursuivre dans cette même voie.

## Structure de gouvernance

Quatre groupes clés permettent d'opérationnaliser le travail du PRC : le comité de direction, les groupes de travail, les intendants des données et l'équipe de recherche.

Le comité de direction est le principal organe décisionnel qui définit les orientations et gère l'ensemble des opérations du partenariat. Chaque partenaire compte au sein du comité un représentant ayant un pouvoir décisionnel pour son organisme. À titre d'exemple, les représentants des conseils scolaires sont les directeurs de l'éducation; celui de McMaster est le registraire de l'Université; et celui de Mohawk est le vice-président des affaires académiques. Le président et le PDG du COQES et de la HCF sont des membres non votants, car aucun des deux organismes ne soumet de données au PRC. Le comité de direction se réunit deux à trois fois par an lors des étapes clés des projets et adhère à un ensemble de principes directeurs qui est décrit dans la section suivante.

Figure 1

Structure de gouvernance du PRC



*Remarque :* Cette figure illustre la structure de gouvernance du PRC : un comité de direction, composé de représentants de chaque établissement partenaire, dirige les groupes de travail désignés, les intendants des données et l'équipe de recherche. Les groupes de travail élaborent des politiques et des processus tandis que les intendants des données, notamment l'analyste des données du partenariat (ADP), et l'équipe de recherche mettent en œuvre les politiques et processus approuvés (c'est-à-dire qu'ils collectent, connectent et analysent les données).

Les groupes de travail sont mis en place par le comité de direction, selon les besoins, afin de proposer des politiques et des processus qui rendent opérationnelles les orientations du comité de direction. Le comité de direction nomme les membres des groupes de travail selon leur affiliation aux partenaires du PRC et en fonction des compétences et de l'expertise nécessaires à chaque étape du projet. Par exemple, le comité de direction a demandé aux chercheurs du COQES et de McMaster d'aider à déterminer les variables pertinentes à inclure dans l'ensemble de données; il a demandé au personnel des technologies de l'information et aux analystes de données des partenaires du secteur de l'éducation d'élaborer des processus de transfert des données; puis il a demandé aux responsables de la protection de la vie privée des établissements d'évaluer les risques et de définir des mesures de sécurité en consultation avec leurs conseillers juridiques.

Le comité de direction a également désigné un intendant de données chez chacun des quatre partenaires du secteur de l'éducation comme responsable de l'exécution des processus d'assemblage et de transfert. Les intendants des données possèdent une connaissance approfondie des données administratives de leur établissement; au cours du projet pilote, on comptait parmi eux des analystes de données de conseils scolaires, un statisticien/programmeur universitaire et un directeur des rapports institutionnels d'un collège. Ensemble, les intendants de données étaient chargés de créer un dictionnaire de données afin de s'assurer que les variables étaient codées de manière uniforme dans tous les établissements, de veiller à ce que les mesures de protection de la vie privée et de sécurité soient respectées dans leurs organismes et de vérifier les données avant leur transfert. Ils ont également répondu aux questions relatives aux données et ont fait part de conseils sur la faisabilité des propositions de recherche.

Le comité de direction a convenu que les données devaient être stockées au HWDSB de façon à en protéger l'accès. Ainsi, l'intendant de données du HWDSB a assumé la responsabilité supplémentaire de fusionner et de dépersonnaliser l'ensemble de données liées. Cette personne était la seule à avoir accès à l'ensemble de données une fois ces dernières assemblées et à effectuer les analyses pour répondre aux questions de recherche déterminées par le comité de direction. En raison de ce rôle unique, le titre d'analyste de données du partenariat (ADP) lui a été attribué. Son rôle consistait également à travailler avec une équipe de recherche, composée de chercheurs du COQES et de McMaster, pour analyser l'ensemble de données et formuler des conclusions de recherche.

Deux accords de partenariat jettent les bases du premier projet (ou phase pilote) et de sa structure de gouvernance. Le premier est un protocole d'entente (PE) entre le COQES et la HCF qui décrit les rôles et responsabilités des cofacilitateurs du PRC. Le second est un accord de partage des données (APD) entre les quatre partenaires du secteur de l'éducation et les cofacilitateurs, qui définit la gouvernance du partenariat ainsi que les protocoles de collecte, de transfert, de gestion et d'utilisation des données pour la première phase de la recherche.<sup>1</sup> Les partenaires ont convenu que les résultats de la phase pilote devaient servir de base à la phase suivante du partenariat et à tous les accords ultérieurs.

## Autorité légale

Les membres du PRC ont entrepris un projet audacieux : en Ontario, les données sont rarement partagées entre les conseils scolaires et les établissements postsecondaires, et encore moins entre plusieurs conseils scolaires et établissements postsecondaires en

---

<sup>1</sup> La recherche de la phase pilote s'est concentrée sur la cohorte d'élèves ayant commencé la 9<sup>e</sup> année à Hamilton pendant l'année scolaire 2010-2011.

même temps (Robson, 2021). Afin de réduire au minimum les préoccupations possibles concernant le partage des données, et conformément aux exigences fixées par le Commissaire à l'information et à la protection de la vie privée de l'Ontario, le comité de direction a créé un groupe de travail chargé de réaliser une évaluation de l'impact sur la protection de la vie privée (EIPVP); cette évaluation visait à recenser les risques pour la vie privée et la sécurité et à prendre des mesures pour les atténuer. L'EIPVP a également servi à préciser l'autorité légale des partenaires en matière de collecte et de partage des données et à établir un processus qui pourrait être partagé et utilisé par d'autres conseils scolaires, collèges et universités.

Le processus d'EIPVP s'est aligné sur le document du Commissaire à l'information et à la protection de la vie privée de l'Ontario intitulé *Planning for Success: Privacy Impact Assessment Guide* (2015) (en anglais seulement). Dans un premier temps, nous avons confirmé que l'EIPVP était nécessaire (c'est-à-dire que nous avons confirmé que le projet impliquerait la collecte, l'utilisation, la conservation, la divulgation, la protection ou l'élimination de renseignements personnels). Le groupe de travail a étudié la *Loi sur l'accès à l'information et la protection de la vie privée* (LAIPVP), la *Loi sur l'accès à l'information municipale et la protection de la vie privée* (LAIMPVP) et la *Loi sur l'éducation* — toutes ces lois sont appelées collectivement les « lois sur la protection de la vie privée » dans la présente section. La LAIPVP et la LAIMPVP décrivent les obligations des établissements publics en matière de collecte, d'utilisation et de divulgation des renseignements personnels. La *Loi sur l'éducation* est la principale loi qui régit les écoles et les conseils scolaires, et la façon dont les services d'éducation sont prodigués dans le système scolaire ontarien financé par les deniers publics » (Commissaire à l'information et à la protection de la vie privée de l'Ontario, 2019). Le groupe de travail sur la protection de la vie privée a déterminé que les lois sur la protection de la vie privée confirment l'autorité légale des partenaires du secteur de l'éducation pour mener à bien ce projet et relier les informations sur les élèves entre les établissements (voir [Annexe A : Autorité légale](#) pour plus de détails sur les articles qui confirment cette autorité).

Essentiellement, les lois sur la protection de la vie privée permettent la divulgation de renseignements personnels à des fins compatibles avec celles pour lesquelles ils ont été recueillis (par exemple, pour éduquer, pour améliorer les résultats scolaires, pour repérer les obstacles et les éliminer), à condition qu'une personne raisonnable s'y attende. Les quatre partenaires du secteur de l'éducation ont acquis les données nécessaires à la création de l'ensemble de données du PRC en toute légalité, pour la bonne administration d'une activité autorisée par la loi (c'est-à-dire à des fins d'éducation des élèves). Le PRC souhaitait utiliser ces données conformément à l'objectif pour lequel elles ont été recueillies (c'est-à-dire à des fins éducatives et pour améliorer les résultats scolaires des élèves). Cela dit, les partenaires hésitaient d'affirmer que les élèves inclus dans l'ensemble



de données, qui commençaient leur 9<sup>e</sup> année au cours de l'année scolaire 2010-2011, ainsi que leurs parents, s'attendaient ou étaient conscients que leurs données allaient être utilisées de cette manière. En fait, leur attente était peut-être même à l'opposé. Les établissements d'enseignement de l'Ontario ont toujours obtenu un consentement pour la recherche au lieu de fixer des attentes et d'expliquer l'autorité légale de recueillir et d'utiliser les renseignements sur les élèves.<sup>2</sup> Ce faisant, les conseils scolaires et, dans une moindre mesure, les collèges et les universités, sous-entendent par inadvertance que la recherche utilisant les données des élèves constitue une atteinte déraisonnable à la vie privée – il s'agit d'une idée fautive et d'un obstacle majeur à la participation à des projets de partage de données comme le PRC.

Conscients de ce problème, les membres du PRC ont adopté une approche conservatrice et ont grandement restreint le contenu de l'ensemble de données, réduisant ainsi le risque que les données ou la recherche qu'elles ont permis de produire puissent être reliées à des individus. Comme nous le décrivons dans les sections suivantes, les partenaires ont développé un processus pour s'assurer que les données étaient aussi « non identifiables » que possible, même au détriment de leur capacité à mettre en évidence des informations riches. Dans les phases à venir, au-delà du projet pilote, les partenaires visent à étendre l'utilité du PRC en analysant les avis de collecte des conseils scolaires et en veillant à ce que les élèves et les parents soient pleinement informés. Au fil du temps, et grâce à une communication efficace de nos premiers résultats de recherche (qui sont résumés dans [Le pouvoir des données connectées](#)), les communautés d'élèves et de parents pourront maintenant s'attendre à ce que les données soient partagées en toute sécurité entre les établissements afin de mieux comprendre les trajectoires d'apprentissage et les obstacles. Cela permettra au PRC de partager et d'analyser des informations détaillées et individualisées en lien avec les données académiques et démographiques afin d'éclairer la prise de décision et d'améliorer les résultats scolaires des élèves de Hamilton.

## Principes directeurs

Le PRC est guidé par un ensemble de principes directeurs, qui sont définis dans un protocole de recherche et reflétés dans l'APD. Le premier de ces principes consiste à mener des recherches dans l'objectif commun de soutenir les élèves dans leurs différents parcours. Il est essentiel d'avoir accès à des données longitudinales pour comprendre comment les décideurs et les praticiens peuvent niveler les chances pour les élèves – en particulier ceux qui ont été historiquement marginalisés et sous-représentés dans

---

<sup>2</sup>Cette affirmation est fondée sur l'expérience professionnelle des trois auteurs, qui ont travaillé avec et pour des conseils scolaires (les auteurs étant : un gestionnaire du service de recherche et d'analyse d'un conseil scolaire, un agent de recherche d'un conseil scolaire et un directeur des partenariats d'un organisme (le COQES) qui travaille fréquemment avec des conseils scolaires).

l'enseignement postsecondaire, notamment les étudiants de première génération (ceux dont les parents ne sont pas titulaires d'un diplôme d'études postsecondaires), les étudiants à faible revenu, les élèves racialisés et les élèves handicapés (Chatoor et coll., 2019, 2022; Ford et coll., 2019; Robson et coll., 2019). Les membres du comité de direction du PRC reconnaissent le rôle important que le partenariat peut jouer afin d'établir les preuves nécessaires pour niveler les chances en matière d'éducation. Au nombre des autres principes directeurs liés à la recherche figurent la nécessité d'approuver d'un commun accord tous les projets de recherche, l'équilibre entre la génération de connaissances et l'action, et la collaboration pour protéger la réputation de tous les partenaires.

En établissant les principes directeurs et les conditions d'un APD, le comité de direction a également reconnu que chaque partenaire du secteur de l'éducation possède son propre comité d'éthique en recherche (CER), avec des règles et des procédures à respecter. Les CER sont des comités indépendants et multidisciplinaires qui veillent à ce que la recherche soit menée de manière éthique, en protégeant les droits et le bien-être des participants à la recherche. Leurs conseils, ainsi que ceux des équipes juridiques des établissements, ajoutent de la valeur et de la sécurité au partenariat.

Afin d'instaurer la confiance entre les parties prenantes (comme les parents et les élèves) et de créer une culture dans laquelle le partage des données est vu comme un outil puissant pour comprendre les lacunes et les obstacles systémiques, le comité de direction a adopté des lignes directrices sur la vie privée et la protection des données, fondées sur l'EIPVP décrite ci-dessus, pour se conformer à la législation pertinente. Ces lignes directrices ont donné lieu à des mesures d'atténuation des risques, qui sont décrites dans la section consacrée à l'assemblage des données, notamment l'application de méthodes de dépersonnalisation et la limitation de l'accès aux données. Le comité de direction a également développé le PRC en utilisant une approche par étapes. La première phase avait pour but de servir de preuve de concept afin de confirmer que les données des partenaires pouvaient être connectées de manière réussie et sûre, en s'appuyant sur les données historiques d'une cohorte d'élèves (c'est-à-dire les élèves ayant commencé la 9<sup>e</sup> année au cours de l'année scolaire 2010-2011).

Les partenaires étaient également préoccupés par les risques liés à la réputation. Pour veiller à ce que la réputation des partenaires ne soit pas entachée par le PRC, l'APD prévoit que la recherche ne peut être utilisée pour classer les organismes ou établir des comparaisons néfastes entre eux. Comme contrepoids, la structure de gouvernance précise que toutes les décisions relatives à la recherche et au fonctionnement doivent être prises par le comité de direction par voie de consensus. Les publications de recherche doivent également être approuvées par le comité de direction.

La liste complète des principes directeurs est présentée à la Figure 2 (et en texte standard de l'[Annexe B : Principes directeurs](#)). Les partenaires ont défini les principes collectivement lors des réunions du comité de direction et du groupe de travail aux premiers stades du partenariat.

Figure 2

Principes directeurs du PRC

## Principes directeurs du PRC



### Recherche

1. Mener des recherches dans l'objectif commun de soutenir les élèves dans leurs différents parcours.
2. Approuver en collaboration tous les projets de recherche.
3. Mener la recherche d'une manière qui intègre et crée un équilibre entre la génération de connaissances et l'action pour le bénéfice mutuel de tous les partenaires.
4. Éviter les questions de recherche ou les projets qui se concentrent sur des établissements particuliers. La recherche ne permettra pas d'identifier des élèves ou des établissements individuels; la recherche ne cherchera pas délibérément à jeter l'opprobre sur un organisme du partenariat, à l'embarrasser ou à lui porter préjudice.
5. Reconnaître et être conscient que chaque partenaire est doté de son propre comité d'éthique de la recherche, chacun ayant ses propres règles et procédures qui doivent être suivies et respectées.



### Gouvernance des données

6. Convenir de lignes directrices en matière de protection de la vie privée et des données, qui respectent les exigences de chaque partenaire, et s'y conformer.
7. Nommer un intendant des données pour chaque partenaire qui sera responsable du respect des droits et des responsabilités en matière de données de son organisme.



### Mobilisation des connaissances

8. Partager les résultats de notre recherche collective avec la communauté dans son ensemble, en permettant à d'autres organismes de voir comment nous travaillons et ce que nous avons appris.
9. Convenir de la manière et du moment où les résultats de la recherche sont communiqués collectivement.
10. Utiliser les résultats de nos recherches pour inspirer des changements politiques qui amélioreront les opportunités et les résultats pour nos élèves.



### Collaboration

11. S'assurer que chaque partenaire ait un rôle et une voix égaux. Les décisions seront prises collectivement; lorsque le contexte ne permet pas d'avoir un accord collectif, les décisions seront négociées autour de notre objectif commun qui est de comprendre le parcours des élèves.
12. Utiliser ce projet comme une occasion d'améliorer la façon dont les données sont recueillies par les partenaires de l'éducation et travailler ensemble, lorsque cela est possible, pour trouver et créer des solutions simplifiées.
13. Déterminer les possibilités pour les élèves de participer à la recherche

*Remarque :* Cette figure montre les quatre principes directeurs du PRC : recherche, gouvernance des données, mobilisation des connaissances et collaboration. Un descriptif textuel standard (c'est-à-dire non visuel) est présenté à l'[Annexe B : Principes directeurs](#).

## Assemblage et stockage des données

L'assemblage de l'ensemble de données du PRC pour la phase pilote a été un processus à plusieurs étapes.<sup>3</sup> Dans un premier temps, les intendants de données des deux conseils scolaires ont regroupé leurs données institutionnelles sur les élèves ayant commencé la 9<sup>e</sup> année au cours de l'année scolaire 2010-2011. L'ADP a fourni une liste de NISO hachés à McMaster et à Mohawk afin que les intendants de données des établissements d'enseignement postsecondaire puissent explorer leurs systèmes d'information sur les étudiants et trouver des données correspondant aux mêmes élèves pour une période allant jusqu'à huit ans (jusqu'à l'automne 2018). Chaque intendant de données s'est assuré que l'ensemble de données qu'il avait créé contenait les variables convenues par les groupes de travail concernés. Les intendants de données ont procédé à une première étape de dépersonnalisation en supprimant tous les identifiants directs, tels que les noms des élèves, et en généralisant<sup>4</sup> les quasi-identifiants sensibles tels que la date de naissance complète et le code postal complet à six chiffres. Pour réduire au minimum le risque que les élèves puissent être identifiés par leur NISO, l'ADP a fourni à chaque intendant de données une formule de hachage<sup>5</sup> pour masquer le NISO dans leurs ensembles de données. Les ensembles de données ainsi obtenus contenant les NISO hachés ont été stockés séparément des bases de données institutionnelles sur les élèves (qui incluent les NISO) afin de réduire davantage le risque de réidentification.

Les intendants de données ont transféré leurs ensembles de données à l'ADP en utilisant un protocole de transfert de fichiers sécurisé (SFTP). Le protocole SFTP permet le transfert sécurisé de données sur un réseau crypté pendant une période déterminée. Pour tester le processus de transfert des données, les intendants de données ont d'abord transféré un échantillon de leurs ensembles de données à l'ADP. Cette opération s'étant avérée concluante, ils ont utilisé le même processus pour transférer les données restantes.

Une fois que les ensembles de données de chaque partenaire du secteur de l'éducation ont été reçus, l'ADP les a reliés en utilisant les NISO hachés. Dans une étape

---

<sup>3</sup> Notre processus est basé sur d'autres modèles, dont le Student Transitions Project de la Colombie-Britannique : une initiative conjointe entre les ministères de l'Éducation et les établissements postsecondaires de la province, qui associe les données de la maternelle à la 12<sup>e</sup> année aux données publiques des établissements postsecondaires afin de mieux comprendre la réussite des étudiants (Colombie-Britannique, n.d.).

<sup>4</sup> La généralisation élimine la granularité des données; ici, les intendants de données n'ont inclus que le mois et l'année pour la date de naissance des étudiants et les trois premières lettres des codes postaux des étudiants.

<sup>5</sup> La formule a été remise en personne par l'ADP aux différents intendants de données sur une clé USB. Les intendants de données étaient les seules personnes de leur établissement respectif à avoir accès à la formule de hachage. Comme mesure de sécurité supplémentaire, les intendants de données ont appliqué du « poivre » (c'est-à-dire une valeur aléatoire qui est ajoutée à la formule de hachage et conservée séparément de la formule).

supplémentaire, l'ADP a effectué un appariement probabiliste<sup>6</sup> en utilisant les variables restantes de quasi-identification (par exemple, le sexe, l'âge calculé). Il a ensuite lié les données sur les demandes d'admission aux études postsecondaires du Service d'admission des collèges de l'Ontario (SACO) et du Centre de demande d'admission aux universités de l'Ontario en utilisant le même processus. L'ensemble de données ainsi obtenu affichait un taux de correspondance de 99,6 %, <sup>7</sup> ce qui signifie que nous avons pu relier avec succès les données des conseils scolaires et des établissements postsecondaires pour presque tous les élèves ayant commencé leur 9<sup>e</sup> année en 2010 à Hamilton et ayant poursuivi leurs études postsecondaires au Collège Mohawk ou à l'Université McMaster. L'ADP a ensuite procédé à une deuxième étape de dépersonnalisation : d'abord, en supprimant le NISO haché et en le remplaçant par un identifiant unique servant aux fins de recherche. Il a ensuite supprimé ou généralisé les variables quasi identifiantes restantes. Une fois les ensembles de données liés, les NISO et la formule de hachage ont été supprimés.<sup>8</sup>

Depuis qu'il a été assemblé, l'ensemble de données a été stocké séparément du système d'information des élèves du HWDSB, derrière plusieurs couches de sécurité, chacune nécessitant une clé. Conformément à l'APD, l'ensemble de données sera conservé par le HWDSB pendant 20 ans, ou pendant la durée du partenariat (selon la période la plus longue). L'ADP a un accès exclusif à l'ensemble de données. Si les membres de l'équipe de recherche souhaitent collaborer aux analyses, ils ne peuvent le faire que sur place avec l'ADP, ou en fournissant des instructions à l'ADP à distance. Aucun autre chercheur ou organisme en dehors du partenariat n'est autorisé à accéder à l'ensemble de données pendant la phase pilote.

---

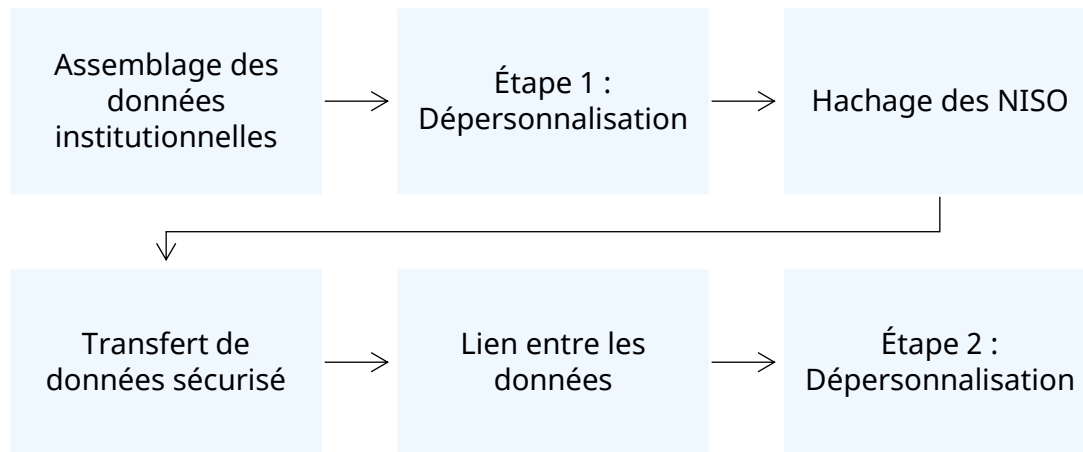
<sup>6</sup> L'appariement probabiliste est une approche statistique utilisée pour mesurer la probabilité que deux entrées dans un ensemble de données représentent le même individu.

<sup>7</sup> Sans identifiants comme les dates de naissance et les noms, nous ne pourrions pas établir de correspondance entre les élèves dont le NISO est incorrect et leur dossier.

<sup>8</sup> La dépersonnalisation de l'ensemble de données final ne devait pas être à toute épreuve, car les données sont conservées sous clé, inaccessibles par Internet et accessibles par une seule personne (l'intendant de données du HWDSB). Pour quiconque connaît le processus de dépersonnalisation, des niveaux de risque peuvent être tolérés en fonction de facteurs contextuels comme le fait que l'ensemble de données sera rendu public ou non, le nombre de personnes qui l'utilisent et les contrôles d'accès. Pour le projet PRC, nous avons utilisé les [lignes directrices sur la dépersonnalisation des données structurées](#) (en anglais seulement) publiées par le Commissaire à l'information et à la protection de la vie privée de l'Ontario.

Figure 3

### Étapes de l'assemblage des données du PRC



*Remarque :* Cette figure est un résumé visuel, sous forme d'organigramme, des étapes de l'assemblage des données du PRC.

## Limites

Même si l'autorisation légale de mener des recherches à des fins éducatives est claire, le comité de direction du PRC a décidé d'adopter une approche conservatrice et d'instaurer la confiance entre les membres et les parties prenantes. Elle a pris les mesures décrites ci-dessus pour réduire au minimum les risques liés à la vie privée et à la sécurité tout en maximisant le potentiel d'amélioration pour les élèves de Hamilton. Pour atteindre cet équilibre, l'ensemble de données que nous avons produit présentait des limites importantes du point de vue de la recherche.

Dans un premier temps, la phase pilote du PRC reposait sur des données historiques. Notre ensemble de données comprenait des informations sur les élèves ayant commencé leur 9<sup>e</sup> année au cours de l'année scolaire 2010-2011. Ce choix a limité les partenaires de recherche à l'examen des variables préexistantes et communes entre eux. Cela impliquait également que nos conclusions ne pouvaient être que rétrospectives.

Les données sont également limitées à une période de 10 ans, ce qui est restrictif; les taux d'obtention de diplôme au niveau secondaire sont conventionnellement calculés en cinq ans, et ceux au niveau postsecondaire en sept ans pour l'université et en trois à sept ans pour le collège, selon le programme. Certains élèves ayant obtenu leur diplôme d'études secondaires après sept ans et qui sont entrés dans un programme d'études postsecondaires de quatre ans ne sont pas considérés comme ayant obtenu leur diplôme d'études collégiales ou universitaires, même s'ils l'ont obtenu dans les années suivantes.

La géographie constitue une troisième limite. L'ensemble de données du PRC nous permet d'explorer l'accès aux études postsecondaires uniquement pour les élèves qui ont fait appel à un service centralisé de demande d'admission (le SACO ou le Centre de demande d'admission aux universités de l'Ontario) pour présenter une demande d'admission et confirmer une acceptation dans un collège ou une université de l'Ontario bénéficiant d'un financement public. Il nous a été impossible de tenir compte des élèves ayant présenté une demande et confirmé une offre d'acceptation dans des établissements situés à l'extérieur de l'Ontario (dans d'autres provinces ou pays). Au-delà des confirmations, nous n'avons pu examiner que les résultats postsecondaires des élèves qui sont restés à Hamilton, soit à McMaster, soit à Mohawk, pour leurs études postsecondaires.

La quatrième limitation importante concerne l'ensemble de données dépersonnalisées, qui repose sur la généralisation et la suppression. Ces deux techniques de dépersonnalisation permettent de préserver la vie privée des individus, mais dans une perspective de recherche, elles limitent le nombre de variables que nous pouvons inclure dans notre analyse et le niveau de nuance que nous pouvons explorer. Par exemple, en raison du faible nombre d'élèves de notre échantillon s'identifiant comme autochtones, cette information a dû être supprimée; de plus, la variable de la langue a dû être généralisée aux catégories « anglais » ou « autre ». En raison de ces protections de la vie privée, la phase pilote de la recherche n'a pu jeter un éclairage sur les parcours des élèves en quête d'équité qui s'identifient comme autochtones ou dont la langue maternelle n'est pas l'anglais, pour ne citer que quelques caractéristiques importantes.

Finalement, nous avons constaté une limite dans le travail avec les données des centres d'application, notamment en ce qui concerne la classification des codes de programme. Statistique Canada utilise une structure de classification qui comprend un total de 10 catégories, réparties entre les sciences, la technologie, le génie et les mathématiques (ou STGM) et les affaires, les sciences humaines, la santé, les arts, les sciences sociales et l'éducation (ou SACHES) (Statistique Canada, 2018). Cette structure aurait été excellente pour trier et analyser les données sur le choix des programmes n'eût été d'un problème : les codes de programme au postsecondaire en Ontario ne sont pas normalisés entre les établissements. Chaque établissement postsecondaire de l'Ontario est libre de catégoriser ses codes de programme comme il l'entend, et il n'existe aucune clé pour faire correspondre ces codes aux codes de la Classification des programmes d'enseignement (CPE) de Statistique Canada. Les chercheurs du COQES ont recodé chaque code de programme qui figurait dans les fichiers de données des demandes d'admission aux études postsecondaires dans l'une des deux grandes catégories de la CEP : STGM ou SACHES. Bien que nécessitant beaucoup de travail, cette solution était la meilleure option disponible. Bien entendu, la classification binaire des codes de



programmes d'études postsecondaires a engendré une perte importante de données, si bien que nous n'avons pas pu approfondir le choix des programmes et les facteurs qui y sont liés.

## Enseignements tirés et prochaines étapes

Les membres du PRC de Hamilton se sont réunis et ont créé un ensemble de données unique permettant de mieux comprendre les facteurs qui influencent l'accès, la persistance et l'achèvement des études postsecondaires au niveau communautaire. Ce faisant, les partenaires ont mis au point un plan directeur qui, nous l'espérons, inspirera et guidera d'autres communautés. Au-delà des enseignements pour les gouvernements et les établissements énoncés dans le rapport [Le pouvoir des données connectées](#), nous avons également tiré deux leçons générales importantes applicables à tous ceux qui s'efforcent d'accomplir ce travail.

### 1. La confiance est essentielle

Les membres du PRC de Hamilton ont cherché à établir des relations de confiance – entre eux et avec les parties prenantes – et à mettre en place des procédures permettant une recherche continue et collaborative en éducation dans leur communauté. En développant et en appliquant les principes et les processus ci-dessus, les membres du PRC ont établi des relations de confiance à travers leurs organismes. Ces relations serviront de base à la poursuite de la recherche et à l'amélioration des politiques et pratiques éducatives à Hamilton.

Les partenaires ont également travaillé à établir la confiance entre les parties prenantes, y compris les élèves et leurs parents – un effort continu qui commence avec l'ouverture, la transparence et la responsabilité. En réponse aux préoccupations des parents concernant l'utilisation des renseignements personnels, les conseils scolaires ont tendance à se montrer prudents dans l'interprétation des lois sur la protection de la vie privée et des décisions du Commissaire à l'information et à la protection de la vie privée. Cette approche conservatrice du traitement des données des élèves peut, par inadvertance, créer des obstacles à la réalisation de recherches robustes dans le domaine de l'éducation. La transparence sur les avantages de la recherche en éducation et les protocoles de sécurité mis en place contribueraient à faire tomber ces barrières.

Les lois sur la protection de la vie privée décrites ci-dessus ouvrent la voie à des collaborations de recherche entre établissements visant à enrichir notre compréhension des trajectoires d'apprentissage. Lorsque les conseils scolaires communiquent clairement leur intention d'utiliser les données dans un but d'amélioration, ils peuvent contribuer à favoriser une culture de la recherche au sein de leurs communautés d'intervenants qui les

accompagnent tout au long de leur parcours éducatif – de la maternelle à la 12<sup>e</sup> année jusqu’au collège ou à l’université (s’ils choisissent l’une de ces voies). Les partenaires du secteur de l’éducation du PRC, grâce à leur participation à ce projet, ont communiqué des avis de collecte qui ouvrent la voie à un continuum de recherche universitaire de l’école publique jusqu’au postsecondaire. Il s’agit d’un changement positif par rapport aux pratiques historiques. Nous encourageons les autres établissements d’enseignement à faire des progrès similaires.

## **2. La gestion de projet et la coordination sont essentielles**

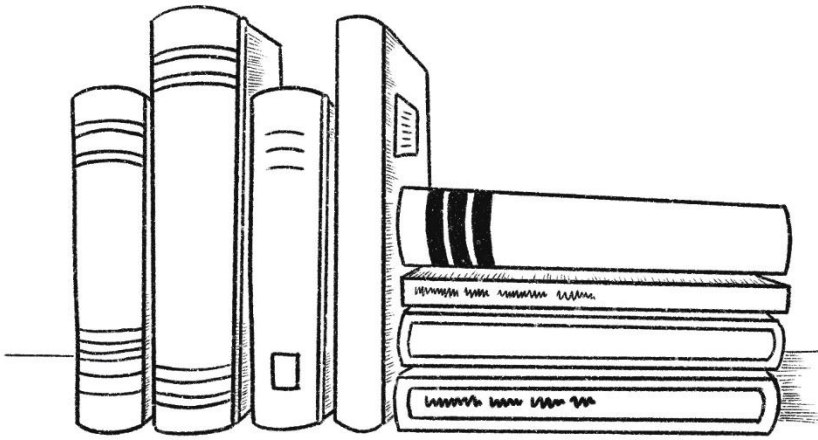
Le PRC était un effort de collaboration reposant sur les principes de la prise de décision partagée. Bien que tous les partenaires aient été essentiels à sa réussite, il convient de souligner la fonction de gestion de projet assumée par les cofacilitateurs du projet. L’intégration du PRC dans les portefeuilles du personnel de la HCF et du COQES était essentielle pour faire avancer le partenariat. Ainsi, le personnel des deux organismes a pu consacrer du temps au soutien des partenaires du secteur de l’éducation, qui se concentraient avant tout sur leur mission principale, à savoir l’éducation des élèves (ce qui était particulièrement difficile pendant la COVID-19). À titre d’exemple, le personnel du COQES et de la HCF a rempli les demandes d’inscription au comité d’éthique de la recherche, organisé les réunions du comité de direction et des groupes de travail, rédigé des documents pour mettre en œuvre les décisions, et répondu aux questions et préoccupations exprimées par chaque partenaire. Nous encourageons les autres communautés à s’assurer qu’au moins un partenaire joue un rôle similaire de gestion et de facilitation du projet.

Nous souhaitons que le partenariat et l’infrastructure de données établis par le projet pilote du PRC puissent servir de base à une recherche collaborative autonome à long terme dans la communauté de Hamilton. Dans les années à venir, alors que les membres du PRC et leurs parties prenantes deviendront de plus en plus à l’aise avec les données connectées, nous prévoyons que les phases futures du PRC auront une portée plus large. Nous nous attendons à ce que les prochaines itérations du PRC englobent plusieurs cohortes d’élèves (y compris le groupe d’élèves le plus récent pour qui nous pouvons étudier les parcours de la 9<sup>e</sup> année à l’école secondaire et leur transition vers le postsecondaire) et incluent des variables supplémentaires et plus nuancées pour l’analyse. Conformément à la législation ontarienne sur la protection de la vie privée, une communication ouverte avec les parties prenantes sur les intentions et les résultats de la recherche permettra aux partenaires de poursuivre l’amélioration de notre ensemble de données. Il sera alors possible de mieux comprendre les facteurs – y compris les obstacles et les aides – qui influent sur les résultats des élèves.

## Références

- Colombie-Britannique, (n.d.). « Student Transitions Project »  
<https://www2.gov.bc.ca/gov/content/education-training/post-secondary-education/data-research/student-transitions-project#:~:text=The%20Student%20Transitions%20Project%20uses,post%2Dsecondary%20education%20and%20graduate>
- Chatoor, K., Mackay, E., et Hudak, L. (2019). *Scolarité des parents et études postsecondaires : La pomme tombe-t-elle loin de l'arbre?* Toronto : Conseil ontarien de la qualité de l'enseignement supérieur <https://heqco.ca/wp-content/uploads/2020/02/FIXED-Formatted-Parental-Ed-f.pdf>
- Chatoor, K., Courts, R., Han, J., Barclay, V., et Colyar, J. (2022) *Programmes Access de l'Ontario : POAEPPI et Passeport pour ma réussite.* Toronto : Conseil ontarien de la qualité de l'enseignement supérieur [https://heqco.ca/wp-content/uploads/2022/01/Access-Programs-in-Ontario\\_FINAL-f.pdf](https://heqco.ca/wp-content/uploads/2022/01/Access-Programs-in-Ontario_FINAL-f.pdf)
- Loi sur l'éducation.* L.R.O (1990). <https://www.ontario.ca/fr/lois/loi/90e02>
- Loi sur l'accès à l'information et la protection de la vie privée.* L.R.O. Chap. F.31. (1990).  
<https://www.ontario.ca/fr/lois/loi/90f31>
- Ford, R., Hui, T. S., et Nguyen, C. (2019). *Participation aux études postsecondaires et revenu des ménages.* Toronto : Conseil ontarien de la qualité de l'enseignement supérieur <https://heqco.ca/wp-content/uploads/2022/02/Formatted-SRDC-PSE-Access-and-Income-f-1.pdf>
- Gallagher-Mackay, K. (2017). *Une infrastructure de données servant à étudier l'accès équitable à l'enseignement postsecondaire en Ontario.* Toronto : Conseil ontarien de la qualité de l'enseignement supérieur <https://heqco.ca/wp-content/uploads/2020/03/FINAL-Data-Infrastructure-f.pdf>
- Commissaire à l'information et à la protection de la vie privée de l'Ontario. (2015). *Planning for Success: Privacy Impact Assessment Guide.* Toronto : Commissaire à l'information et à la protection de la vie privée de l'Ontario.  
<https://www.ipc.on.ca/wp-content/uploads/2015/05/planning-for-success-pia-guide.pdf>
- Commissaire à l'information et à la protection de la vie privée de l'Ontario. (2016). *De-identification Guidelines for Structured Data.* Toronto : Commissaire à l'information et à la protection de la vie privée de l'Ontario. <https://www.ipc.on.ca/wp-content/uploads/2016/08/Deidentification-Guidelines-for-Structured-Data.pdf>

- Commissaire à l'information et à la protection de la vie privée de l'Ontario. (2019). *Guide sur la protection de la vie privée et l'accès à l'information dans les écoles de l'Ontario*. Toronto : Commissaire à l'information et à la protection de la vie privée de l'Ontario. <https://www.cipvp.ca/wp-content/uploads/2019/01/guide-to-privacy-access-in-ont-schools-FRE.pdf>
- Loi sur l'accès à l'information municipale et la protection de la vie privée*. L.R.O. chap. M. 56. (1990). <https://www.ontario.ca/fr/lois/loi/90m56>
- Robson, K., Maier, R., Anisef, P., et Brown, R. S. (2019). *La réussite au secondaire et l'accès aux études postsecondaires*. Toronto : Conseil ontarien de la qualité de l'enseignement supérieur <https://heqco.ca/wp-content/uploads/2020/02/Formatted-CRP-report-f.pdf>
- Robson, K. (2021). An Essay on the Challenges of Doing Education Research in Canada. *Journal of Applied Social Science*, 15(2),183–196. DOI :10.1177/19367244211003471 <https://journals.sagepub.com/doi/full/10.1177/19367244211003471>
- Statistique Canada. (2018). *Variante de la CPE 2016 - Regroupements STGM et SACHES*. Classifications statistiques. [https://www23.statcan.gc.ca/imdb/p3VD\\_f.pl?Function=getVD&TVD=401856](https://www23.statcan.gc.ca/imdb/p3VD_f.pl?Function=getVD&TVD=401856)



# **Plan directeur du PRC : Comment nous avons construit une infrastructure communautaire de données**

Annexes

## Annexe A : Autorité légale

Selon le groupe de travail sur la protection de la vie privée, l'autorité légale pour mener à bien ce projet et relier les informations sur les élèves entre les établissements est confirmée par ce qui suit :

1. L'alinéa 266.3(3) de la *Loi sur l'éducation* précise que « [l]e ministre et toute personne ou entité prescrite peuvent recueillir, utiliser ou divulguer des numéros d'immatriculation scolaire de l'Ontario, ou en exiger la production, à des fins liées à l'administration, au financement ou à la planification de l'éducation ou à la recherche dans ce domaine. » Cela confirme notre utilisation des NISO comme clé principale pour mener des recherches en éducation de la manière que nous avons décrite.
2. L'alinéa 14(1)(e) de la LAIMPVP et l'alinéa 21(1)(e) de la LAIPVP prévoient qu'un établissement peut divulguer des renseignements personnels :
  - « e) à des fins de recherche si :
    - (i) la divulgation est conforme aux conditions ou à l'utilisation envisagées au moment où ces renseignements ont été divulgués, recueillis ou obtenus,
    - (ii) les fins de recherche à l'origine de la divulgation ne peuvent être raisonnablement atteintes que si les renseignements sont divulgués sous une forme qui permette l'identification individuelle,
    - (iii) la personne devant recevoir le document a accepté de se conformer aux conditions relatives à la sécurité et au caractère confidentiel qui sont prescrites par les règlements. »
3. L'article 10 du Règlement 460, pris en application de la LAIPVP, définit les exigences en matière de sécurité et de confidentialité qui doivent être convenues avant que l'établissement puisse divulguer des renseignements personnels à des fins de recherche. Ces dispositions garantissent le respect de la clause (iii), ci-dessus :
  1. La personne n'utilise les renseignements qu'à des fins de recherche précisées dans l'accord ou pour lesquelles elle a reçu l'autorisation écrite de l'institution.
  2. La personne nomme dans l'accord les autres personnes à qui sera accordé l'accès aux renseignements personnels sous une forme dans laquelle le particulier concerné par ces renseignements peut être identifié.

3. Avant de divulguer les renseignements personnels aux autres personnes visées à la disposition 2, la personne conclut un accord avec celles-ci pour veiller à ce qu'elles ne les divulguent pas à d'autres personnes.
4. La personne conserve les renseignements dans un endroit sûr dont l'accès n'est accordé qu'à la personne et aux personnes à qui l'accès est accordé aux termes de la disposition 2.
5. La personne détruit tous les identificateurs individuels contenus dans les renseignements au plus tard à la date précisée dans l'accord.
6. La personne ne communique avec aucun particulier concerné par ces renseignements personnels, directement ou indirectement, sans obtenir au préalable l'autorisation écrite de l'institution.
7. La personne veille à ce qu'aucun renseignement personnel ne soit utilisé ou divulgué sous une forme dans laquelle le particulier concerné par ce renseignement peut être identifié, à moins d'obtenir l'autorisation écrite de l'institution.
8. La personne avise l'institution par écrit immédiatement si elle apprend que les conditions énoncées au présent article n'ont pas été observées.

## Annexe B : Principes directeurs

### *Recherche*

1. Mener des recherches dans l'objectif commun de soutenir les élèves dans leurs différents parcours.
2. Approuver en collaboration tous les projets de recherche.
3. Mener la recherche d'une manière qui intègre et crée un équilibre entre la génération de connaissances et l'action pour le bénéfice mutuel de tous les partenaires.
4. Éviter les questions de recherche ou les projets qui se concentrent sur des établissements particuliers. La recherche ne permettra pas d'identifier des élèves ou des établissements individuels; la recherche ne cherchera pas délibérément à jeter l'opprobre sur un organisme du partenariat, à l'embarrasser ou à lui porter préjudice.
5. Reconnaître et être conscient que chaque partenaire est doté de son propre comité d'éthique de la recherche, chacun ayant ses propres règles et procédures qui doivent être suivies et respectées.

### *Gouvernance des données*

6. Convenir de lignes directrices en matière de protection de la vie privée et des données, qui respectent les exigences de chaque partenaire, et s'y conformer.
7. Nommer un intendant des données pour chaque partenaire qui sera responsable du respect des droits et des responsabilités en matière de données de son organisme.

### *Mobilisation des connaissances*

8. Partager les résultats de notre recherche collective avec la communauté dans son ensemble, en permettant à d'autres organismes de voir comment nous travaillons et ce que nous avons appris.
9. Convenir de la manière et du moment où les résultats de la recherche sont communiqués collectivement.
10. Utiliser les résultats de nos recherches pour inspirer des changements politiques qui amélioreront les opportunités et les résultats pour nos élèves.

### *Collaboration*

11. S'assurer que chaque partenaire ait un rôle et une voix égaux. Les décisions seront prises collectivement; lorsque le contexte ne permet pas d'avoir un accord collectif, les décisions seront négociées autour de notre objectif commun qui est de comprendre le parcours des élèves.



- 12.** Utiliser ce projet comme une occasion d'améliorer la façon dont les données sont recueillies par les partenaires de l'éducation et travailler ensemble, lorsque cela est possible, pour trouver et créer des solutions simplifiées.
- 13.** Déterminer les possibilités pour les élèves de participer à la recherche.