



Système d'information amélioré sur les étudiants (SIAE): vue d'ensemble

Centre de la statistique de l'éducation
Statistique Canada

Août 2002

Table des matières

1	INTRODUCTION.....	2
2	HISTORIQUE : GENÈSE DU SIAE.....	3
3	CONTENU DE LA BASE DE DONNÉES DU SIAE.....	4
4	DONNÉES SUR LES ÉTABLISSEMENT ET LES PROGRAMMES.....	5
5	COMMENT LE SIAE AMÉLIORERA L'INFORMATION SUR LES ÉTABLISSEMENTS ET LES PROGRAMMES.....	6
6	DONNÉES SUR LES ÉTUDIANTS.....	7
7	COMMENT LE SIAE AMÉLIORERA L'INFORMATION SUR LES ÉTUDIANTS.....	8
8	LA POSSIBILITÉ DE VALEUR AJOUTÉE PAR LE COUPLAGE DE L'INFORMATION AVEC D'AUTRES SOURCES DE DONNÉES.....	10
9	LA CONFIDENTIALITÉ, LA PROTECTION DE LA VIE PRIVÉE ET LA SÉCURITÉ DES RENSEIGNEMENTS.....	13
	LE POUVOIR LÉGAL SE RATTACHANT AU SIAE	14
	LA CONFIDENTIALITÉ DES RENSEIGNEMENTS.....	15
	LA SÉCURITÉ DE LA TRANSMISSION DES DONNÉES DU SIAE	15
	LA SÉCURITÉ DES RENSEIGNEMENTS.....	16
	LA CONFIDENTIALITÉ DES PRODUITS STATISTIQUES.....	17
	LE COUPLAGE DES ENREGISTREMENTS.....	17
	L'AVIS AUX ÉTUDIANTS DE L'UTILISATION QUI EST FAITE DES RENSEIGNEMENTS LES CONCERNANT.....	18
	LE NUMÉRO NATIONAL D'ÉTUDIANT DU SIAE (NNE-SIAE)	19
	LE NUMÉRO D'ASSURANCE SOCIALE (NAS).....	19
	LA CONSERVATION DES ENREGISTREMENTS.....	20
	L'ACCÈS DES PARTICULIERS AUX ENREGISTREMENTS QUI LES CONCERNENT.....	21
	L'ACCÈS AUX DONNÉES DU SIAE AUX FINS D'ANALYSE.....	21
10	PRODUITS ET SERVICES.....	22
	ANNEXE.....	24
	ÉBAUCHE DE TEXTE À INCLURE DANS LES FORMULES D'INSCRIPTION ET LES ANNUAIRES.....	24

1 INTRODUCTION

Statistique Canada met au point le Système d'information amélioré sur les étudiants sous l'égide du Conseil des statistiques canadiennes de l'éducation. Le Conseil est un organisme conjoint de Statistique Canada et des sous-ministres provinciaux et territoriaux responsables de l'éducation. Il est coprésidé par le statisticien en chef du Canada et un sous-ministre. Les ministères agissent collectivement par l'entremise du Conseil des ministres de l'Éducation (Canada).

Le premier objectif du SIAE est de répondre aux besoins stratégiques et de planification dans le domaine des études postsecondaires et de la transition vers le marché du travail.

La place du Canada dans l'économie mondiale dépend en grande partie des connaissances et des compétences de ses citoyens. Le système d'enseignement postsecondaire est un outil important qui permet de dispenser les connaissances et les compétences requises. Pour que le Canada puisse rivaliser dans la nouvelle économie, il est essentiel que les étudiants puissent faire des choix de programmes éclairés et que l'on puisse mobiliser efficacement les ressources en éducation.

Le présent document offre un aperçu du SIAE et de la nature des renseignements qu'il contient. Il aborde des questions telles que la protection de la vie privée, la confidentialité et la sécurité des données. Ces questions sont également examinées plus en détail dans un document connexe intitulé *Système d'information amélioré sur les étudiants : réponses aux inquiétudes relatives à la protection de la vie privée et des renseignements personnels*.

Le SIAE est conçu comme un inventaire complet de tous les établissements postsecondaires canadiens et des programmes et cours qu'ils offrent, de même que des renseignements sur les caractéristiques démographiques, sur les programmes et sur les cours de chaque étudiant inscrit dans ces établissements.

Le SIAE remplace les enquêtes administratives actuelles sur les effectifs et sur les diplômés postsecondaires. Le Système d'information statistique sur la clientèle universitaire (SISCU), le Système d'information statistique sur la clientèle des collèges communautaires (SISCCC) et l'Enquête sur les programmes de formation professionnelle (EPFP) sont en place depuis bien des années et avaient besoin d'être mis à jour. Le SIAE a été mis au point pour s'attaquer aux points faibles des enquêtes existantes et pour donner plus de renseignements sur les résultats des étudiants afin de répondre aux besoins stratégiques et de planification du Conseil des statistiques canadiennes de l'éducation et d'autres intervenants en éducation.

2 HISTORIQUE : GENÈSE DU SIAE

Statistique Canada a commencé à compiler et à publier des données sur les effectifs et les diplômés postsecondaires dans les années 70. Depuis le début des années 70, en vertu des pouvoirs de la *Loi sur la statistique*, ces données sont tirées des dossiers des étudiants soumis par tous les collèges et universités du Canada. Elles proviennent des systèmes administratifs des établissements postsecondaires. Dans certains cas, ces établissements soumettent directement leurs données à Statistique Canada. Dans d'autres, ils les remettent à des ministères ou organismes provinciaux qui, à leur tour, en rendent compte à Statistique Canada.

Au fil des ans, ces données nous ont permis de mieux comprendre le système d'enseignement postsecondaire et revêtent une importance tant nationale qu'internationale. À l'échelle nationale, les données servent à faire des comparaisons et à établir des tendances annuelles au niveau des provinces. Combinées à l'information sur les ressources humaines et financières consacrées à ce niveau d'éducation, ces données apportent des renseignements essentiels sur l'efficacité et l'efficacé du système d'enseignement postsecondaire. À l'échelle internationale, les données situent le Canada dans les comparaisons qu'effectuent l'Organisation de coopération et de développement économiques (OCDE) et l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO). Ces données sont de plus en plus capitales pour la position du Canada dans les discussions sur le libre-échange.

Ces dernières années cependant, le public, des chercheurs, les ministères provinciaux et les établissements d'enseignement postsecondaire ont demandé des données plus abondantes et améliorées pour s'attarder à des questions toujours plus complexes. En 1997, le Conseil des ministres de l'Éducation (Canada), le CMEC, et le Centre de la statistique de l'éducation se sont lancés ensemble dans un exercice de planification stratégique qui consistait à définir les questions stratégiques et de gestion liées à l'éducation et l'information nécessaire pour y répondre. Tous les ministères provinciaux et territoriaux de l'éducation et plus d'une centaine d'associations nationales et provinciales ont été consultés. Le SIAE a été intégré dans le plan stratégique élaboré à cette fin pour le Centre de la statistique de l'éducation. Ce plan a été approuvé en 1997 par le Conseil des statistiques canadiennes de l'éducation (CSCE).

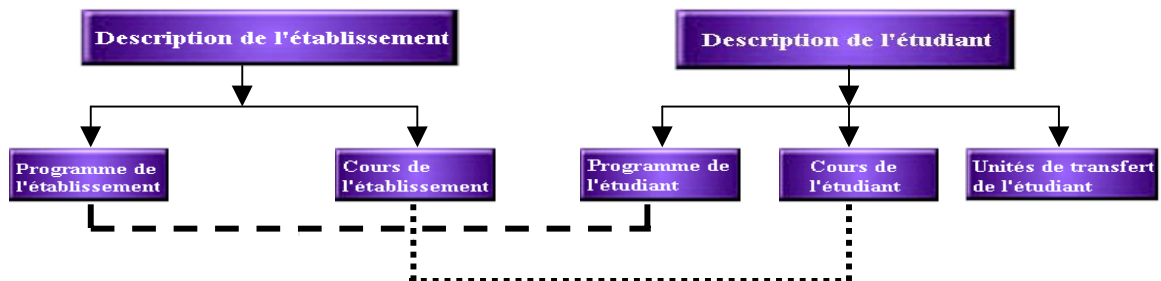
Statistique Canada a assuré le financement initial du projet du SIAE. Deux projets pilotes, l'un réalisé par des universités des Maritimes et l'autre, avec des collèges de l'Ontario, ont servi à définir les besoins en données et à évaluer la faisabilité du SIAE. Des consultations ont aussi eu lieu auprès des gouvernements et des établissements qui ont apporté des données afin de peaufiner davantage les besoins en données. En raison des résultats positifs des projets pilotes et des consultations, l'Initiative de la recherche sur les politiques

(IRP) du gouvernement fédéral a accordé des fonds pour la mise en œuvre. L'IRP a pour objectif d'appuyer les recherches qui permettent d'élaborer des politiques et des initiatives destinées à améliorer la position concurrentielle du Canada dans l'économie mondiale. L'IRP reconnaît le rôle que le SIAE pourrait jouer en accroissant l'efficacité du système d'enseignement postsecondaire au Canada. Dernièrement, d'autres ministères fédéraux ont montré de l'intérêt pour les travaux. Développement des ressources humaines Canada (DRHC) est prêt à faire un apport financier parce que le SIAE permettra d'offrir de meilleures données de planification et de dresser une liste consolidée, complète et exacte des établissements postsecondaires et des programmes qu'ils offrent. Cet examen complet et le registre exact des établissements et des programmes sont essentiels au travail du Programme canadien de prêts aux étudiants géré par DRHC.

3 CONTENU DE LA BASE DE DONNÉES DU SIAE

La base de données du SIAE contient des données à deux niveaux : (i) des données sur les établissements, les programmes et les cours et (ii) des données sur les étudiants. Les données sur les établissements, programmes et les cours comprend la liste complète de tous les établissements postsecondaires au Canada et un inventaire complet de tous les programmes et cours qui y sont offerts. Les données sur les étudiants renferment des renseignements sur les caractéristiques démographiques, sur les programmes et sur les cours des étudiants inscrits dans ces établissements. Comme l'illustrent les chiffres ci-joints et tel que décrit dans les parties 4 et 5, la base de données du SIAE est alimentée par sept fichiers d'entrée décrivant l'étudiant ou l'établissement et les programmes offerts.

Fichiers d'entrées du SIAE



Les données du SIAE sont déclarées une fois l'an, soit directement par les établissements, soit indirectement par les organismes de coordination qui en

rendront compte en leur nom. Le rapport final des données du SIAE, sur toute l'année scolaire, doit être présenté en juillet. Les données finales donneront une image complète de l'activité scolaire sur une période complète de 12 mois.

4 DONNÉES SUR LES ÉTABLISSEMENT ET LES PROGRAMMES

Trois fichiers décrivent les programmes et les cours d'un établissement :

Fichier Description de l'établissement : Énumère et décrit les différentes périodes d'activité scolaire suivant lesquelles les cours et les programmes sont organisés.

Fichier Programme de l'établissement : Donne le titre du programme, sa durée, l'attestation d'études à laquelle il mène, ses préalables et d'autres caractéristiques de chacun des programmes offerts par l'établissement.

Fichier Cours de l'établissement : Donne le titre du cours, sa durée, sa valeur en unités et d'autres caractéristiques de chacun des cours offerts par l'établissement.

Ces fichiers regroupent les renseignements qui sont déjà du domaine public, en format électronique ou sur papier. Le nom des établissements et les programmes qu'ils offrent figurent dans les publications de chaque établissement, dans les publications du gouvernement produites aux fins du counselling des étudiants, dans les publications du secteur privé ou dans les publications d'organismes provinciaux ou nationaux, comme l'Association des universités et collèges du Canada (AUCC) et l'Association des collèges communautaires du Canada (ACCC). Toutefois, les bases de données existantes ont toutes des limites en ce qui touche l'intégralité, le champ d'application et le manque de codage standard de l'information sur les programmes. Le SIAE dépassera ces limites en regroupant l'information en une seule source harmonisée pour tous les systèmes d'enseignement postsecondaire.

Ce registre complet et exact des établissements et des programmes est essentiel au travail de Statistique Canada en planification et en exécution du programme statistique sur l'éducation postsecondaire.

5 COMMENT LE SIAE AMÉLIORERA L'INFORMATION SUR LES ÉTABLISSEMENTS ET LES PROGRAMMES

- a) Le SIAE donnera des renseignements complets et exacts sur les programmes offerts par les établissements postsecondaires, codés dans la Classification des programmes d'enseignement (CPE). Actuellement, il n'existe aucune base de données aussi complète sur les établissements postsecondaires ni d'inventaire codé des programmes. Les listes des établissements que tient actuellement le Centre de la statistique de l'éducation sont incomplètes et des comparaisons ne peuvent être faites entre les systèmes employés pour classer les programmes d'études selon les divers paliers d'enseignement postsecondaire, les universités par rapport aux collèges et aux établissements de formation professionnelle et technique. Le SIAE permettra de diffuser des renseignements comparables sur les programmes par niveau d'instruction.
- b) Grâce à la CPE, le SIAE permettra de faire des comparaisons à l'échelle internationale. La CPE a été acceptée comme norme statistique de classification des programmes par le National Center for Education Statistics du ministère de l'Éducation des États-Unis et par Statistique Canada.
- c) La base de données sur les établissements et les programmes du SIAE renfermera des données de la base de sondage de l'Enquête nationale auprès des diplômés (END). Pour concevoir l'échantillon de l'END, il faut dresser la liste des établissements et du nombre de diplômés par programme. Dans les END antérieures, le manque de comparabilité entre les données des programmes conservées par Statistique Canada pour les divers paliers d'enseignement postsecondaire posait un problème.
- d) Sous réserve des accords portant sur la communication des renseignements² prévus dans la *Loi sur la statistique*, Statistique Canada pourrait échanger de l'information sur les établissements et les programmes avec d'autres

² L'article 12 de la *Loi sur la statistique* contient des dispositions précises sur les accords de communication des renseignements :

(Statistique Canada) peut conclure avec tout ministère ou toute municipalité ou toute personne morale un accord portant sur la communication des renseignements recueillis de l'intéressé par Statistique Canada ou ce ministère ou cette personne morale pour leur compte ainsi que sur les classifications ou publications subséquentes fondées sur ces renseignements.

Un accord conclu en vertu du paragraphe 1 prévoit :

- a) que l'intéressé est informé, par avis à cet effet, que les renseignements sont recueillis pour le compte de Statistique Canada et du ministère ou de la personne morale en cause, selon le cas;
- b) que lorsque l'intéressé donne par écrit au statisticien en chef avis de son opposition à la communication des renseignements par Statistique Canada, ceux-ci ne peuvent être communiqués au ministère ou à la personne morale...

organismes qui auraient besoin de ces renseignements. Ceci réduirait le fardeau du répondant imposé aux établissements qui doivent actuellement donner ce genre de renseignements à de nombreux organismes; cela limiterait l'inefficacité de dresser des listes distinctes; et cela réglerait les problèmes de qualité des listes existantes.

6 DONNÉES SUR LES ÉTUDIANTS

Quatre dossiers donnent des renseignements descriptifs sur les étudiants et chaque programme et cours auxquels ils étaient inscrits pendant le cycle de déclaration. Contrairement aux fichiers sur les établissements et les programmes, ces fichiers contiennent des données sur chaque étudiant.

Fichier Description de l'étudiant : Nom de l'étudiant, date de naissance, sexe, numéro d'assurance sociale (NAS), coordonnées personnelles, études antérieures et caractéristiques, comme le fait que l'étudiant s'identifie volontairement comme membre d'une minorité visible, Autochtone ou personne handicapée. Il y a un enregistrement par étudiant et par établissement. Ce fichier renferme aussi le numéro national d'étudiant du SIAE (NNE-SIAE). Le NNE est un identificateur unique qui sert à coupler les enregistrements reçus à ceux qui se trouvent déjà dans la base de données nationale. Le NNE-SIAE est décrit plus loin dans le présent document.

Fichier Programme de l'étudiant : Renferme un enregistrement pour chaque programme auquel l'étudiant était inscrit au cours du cycle de déclaration. L'enregistrement du programme de l'étudiant comprend la date de début/fin du programme de l'étudiant, sa situation (encore inscrit, a terminé le programme/obtenu son diplôme, s'en est retiré, n'a pas réussi à l'achever, autre), ses majeures, ses unités de transfert, les droits qui lui ont été facturés, les unités qu'il a acquises jusque-là et d'autres caractéristiques du programme de l'étudiant enregistrées par l'établissement.

Fichier Cours de l'étudiant : Renferme un enregistrement pour chaque cours auquel l'étudiant était inscrit pendant le cycle de déclaration. Cet enregistrement inclut de l'information sur la date de début/fin du cours de l'étudiant, sa situation (encore inscrit, a terminé le cours/obtenu son diplôme, s'en est retiré, n'a pas réussi à l'achever, s'il est inscrit comme auditeur libre, autre), les unités qu'il a acquises, les droits qui lui ont été facturés et d'autres caractéristiques du cours de l'étudiant enregistrées par l'établissement.

Fichier Unités de transfert de l'étudiant : Renferme un enregistrement pour chaque cours à la suite duquel l'étudiant s'est vu accorder une unité à partir de cours et/ou d'expériences d'apprentissage suivis à l'extérieur de l'établissement fréquenté. Dans la mesure où cette information est conservée dans le système

administratif de l'établissement, l'enregistrement sur le cours menant à une unité de transfert renferme le nombre d'unités de transfert, la source des unités et la date où elles ont été accordées.

7 COMMENT LE SIAE AMÉLIORERA L'INFORMATION SUR LES ÉTUDIANTS

Les données sur les étudiants du SIAE remplaceront les enquêtes administratives actuelles sur les effectifs et sur les diplômes postsecondaires : le Système d'information statistique sur la clientèle universitaire (SISCU), le Système d'information statistique sur la clientèle des collèges communautaires (SISCCC) et l'Enquête sur les programmes de formation professionnelle (EPFP). Voici des exemples des manières dont le SIAE améliorera l'information répondant à ces besoins stratégiques :

- a) Les enquêtes actuelles auprès des collèges et des universités s'appellent des enquêtes « instantanées ». Les chiffres sur les effectifs sont établis à un moment donné dans l'année. Il manque donc dans les statistiques courantes un nombre important de renseignements sur les effectifs. Le SIAE, en saisissant toute l'activité scolaire de l'année sur une période de 12 mois, permettra d'avoir une compréhension complète de la question.
- b) Les enquêtes actuelles ne saisissent que le domaine d'études principal. En comparaison, le SIAE montrera comment les programmes et l'amalgame de cours d'un étudiant évoluent en réaction aux exigences du marché du travail et aux autres demandes.
- c) Le SIAE recueillera des données sur les inscriptions au niveau des cours, notamment des renseignements sur la durée et la valeur des unités des différents cours, de sorte qu'on pourra, de façon standard et exacte, mesurer la charge de travail des étudiants afin de remplacer la mesure imprécise actuelle connue sous le nom d'« équivalent temps plein » ou « ETP ». C'est une variable qui exige une mesure précise. Elle se répercute sur la durée et le coût des études, sur la persévérance et sur la capacité du système d'enseignement postsecondaire.
- d) Les enquêtes actuelles ne donnent pas d'information sur les étudiants étrangers qui font leurs études par l'intermédiaire d'établissements d'enseignement canadiens, que ce soit sur des campus au Canada ou outre-mer, à domicile ou en milieu de travail par l'entremise de l'éducation à distance. Le SIAE apportera des données sur tous ces étudiants. Dans l'environnement mondial, les gouvernements désirent de plus en plus appuyer et évaluer les répercussions de l'expérience d'apprentissage suivie à l'extérieur du Canada.

- e) Le SIAE renferme de nouveaux éléments de données, comme les préalables des programmes, l'alternance travail-études, les bourses d'études et de subsistance, les cours présentés en vertu de contrats, la capacité d'admission à un programme et le mode de présentation d'un cours. Par ailleurs, le SIAE a été conçu pour tenir compte des changements ultérieurs apportés à la structure de l'enseignement postsecondaire et de l'émergence des nouveaux besoins en information.
- f) Acquisition continue du savoir - Le SIAE donnera de l'information sur la participation subséquente à des études postsecondaires. Combiné à des enquêtes-échantillons et à d'autres bases de données, il permettra de mesurer les répercussions de l'apprentissage continu sur la carrière et sur le revenu (la productivité).
- g) Efficacité du système d'éducation - Les gouvernements se préoccupent des résultats du système d'éducation par rapport aux ressources qu'ils y consacrent. Des mesures, comme le nombre de personnes qui terminent leurs études, le temps qu'elles prennent pour les terminer, leur persévérance et leur taux de rétention, donnent des indications sur le lien entre l'investissement dans l'éducation et les résultats obtenus. Le SIAE améliorera la capacité à standardiser ces mesures et à établir des comparaisons.
- h) Transférabilité des unités - De plus en plus d'étudiants passent aujourd'hui d'un domaine, d'un établissement et d'un niveau d'études à un autre (p. ex., entre des collèges et des universités) afin d'acquérir le bagage de connaissances, de compétences et d'attestations d'études qu'ils souhaitent. Le nombre d'unités de transfert accordées par les établissements a de fortes répercussions sur l'utilisation efficace de l'argent investi en éducation. Le SIAE apportera des mesures comparatives sur l'octroi d'unités de transfert.
- i) Réalisation des programmes coopératifs - Beaucoup d'établissements, souvent des collèges et des universités, offrent aujourd'hui une gamme de programmes suivant une diversité de formules dites articulées, coopératives et à courtage. Des programmes peuvent, par exemple, commencer par les soins infirmiers au niveau collégial et se poursuivre par les sciences infirmières au niveau universitaire, où les programmes collégiaux sont pleinement reconnus. Les programmes coopératifs peuvent combiner des cours du niveau collégial à des cours du niveau universitaire dans un domaine d'études appliquées. Dans les programmes à courtage, les établissements prennent des dispositions pour faire suivre à leurs étudiants certains de leurs cours dans un deuxième ou un troisième établissement. Le SIAE servira de mesure de l'augmentation du nombre de tels programmes offerts.
- j) Partenariats avec l'industrie - Beaucoup d'établissements concluent aujourd'hui des partenariats avec le secteur privé pour offrir des services

d'enseignement, comme des cours/programmes menant à des grades sur mesure et des cours/programmes de formation des employés. Le SIAE offrira une mesure de l'ampleur et de la nature de telles activités.

- k) Mobilité des étudiants - Les gouvernements provinciaux s'intéressent à la mobilité des étudiants. Certaines provinces enregistrent une importante migration de sortie des étudiants, tandis que l'inverse est vrai pour d'autres. Le SIAE permettra d'étudier les migrations d'entrée et de sortie des étudiants.
- l) Parcours des étudiants - Le SIAE suivra le parcours des étudiants au fur et à mesure qu'ils passent d'un domaine, d'un établissement d'enseignement et d'un collège (ou d'une université) à un autre et du statut d'étudiant à plein temps à celui d'étudiant à temps partiel.
- m) Progression de la population des écoles secondaires - Les provinces s'intéressent aussi à la progression future des élèves de leur réseau d'écoles secondaires. Le SIAE apportera aux provinces de l'information sur le pourcentage d'élèves qui font des études postsecondaires au Canada, sur le niveau de ces études et sur leurs résultats.

8 LA POSSIBILITÉ DE VALEUR AJOUTÉE PAR LE COUPLAGE DE L'INFORMATION AVEC D'AUTRES SOURCES DE DONNÉES

Le couplage de l'information obtenue des établissements à d'autres sources de statistiques au microniveau permettra de répondre à des besoins stratégiques et de planification importants. À l'avenir, on demandera aux répondants des enquêtes-échantillons volontaires, comme l'Enquête auprès des jeunes en transition, la permission de combiner les données qu'ils apporteront aux données administratives que renfermera le SIAE.

Note sur le couplage des enregistrements

Statistique Canada recueille des renseignements de diverses sources. Le Centre de la statistique de l'éducation possède des données provenant d'enquêtes-échantillons, comme l'Enquête auprès des jeunes en transition et l'Enquête nationale auprès des diplômés, de même que des données d'enquêtes administratives, comme le SIAE.

Il est possible, sous réserve de conditions extrêmement contrôlées et uniquement à des fins de recherche ou de statistiques, de coupler deux ou plusieurs enquêtes à Statistique Canada. Dans ce cas, le « couplage » se limite à une période déterminée et à des fins d'analyse très particulières.

Les contrôles stricts exercés sur le couplage des enregistrements figurent dans la politique sur le couplage des enregistrements de Statistique Canada. Le statisticien en chef doit approuver toutes les propositions de couplage d'enregistrement et une description de tous les fichiers couplés figurera dans le site Web de Statistique Canada.

De même, il sera possible d'ajouter au SIAE des renseignements sur les prêts aux étudiants ou sur leur revenu après l'obtention du diplôme, grâce à des couplages à des fichiers administratifs. Il est à noter que tout couplage doit être auparavant approuvé en vertu de la politique sur le couplage des enregistrements de Statistique Canada. Les fichiers couplés serviront uniquement d'échantillons, sans identificateur personnel, utilisés uniquement à des fins de recherche et de statistiques approuvées et seront conservés uniquement pendant la durée de ces études.

L'utilisation du NAS comme identificateur commun et unique entre le SIAE et d'autres fichiers couplés faciliterait beaucoup la stratégie de couplage des échantillons décrite ci-dessus. Cette stratégie pourrait comprendre la sélection d'un échantillon d'enregistrements dans le SIAE qui soit représentatif de divers programmes et niveaux d'études dans les provinces ou les territoires. Le couplage pourrait se faire uniquement entre ces enregistrements dans les autres fichiers qui seraient appariés au NAS de l'échantillon des enregistrements du SIAE. Sans identificateur unique, on ne pourrait pas employer de telles stratégies d'échantillonnage en utilisant des échantillons des deux ensembles de données administratives.

Voici des exemples des avantages du couplage des données du SIAE à d'autres sources d'information :

- a) Enquête auprès des jeunes en transition (EJET) :
La combinaison des données du SIAE et de l'EJET permettrait de comprendre l'accès aux études postsecondaires, le taux de rétention, la persévérance, le temps mis pour terminer les études et les résultats à la lumière des caractéristiques des jeunes, de leur comportement, de leur attitude et de leurs aspirations, de leurs antécédents familiaux et de leur expérience au cours de leurs études secondaires et postsecondaires.
- b) Enquête nationale auprès des diplômés (END) :
La combinaison du SIAE et de l'END donnerait de l'information beaucoup plus détaillée sur les programmes et les cours que ne le permet actuellement l'END, ce qui aiderait à comprendre les connaissances et les compétences acquises par rapport aux résultats sur le marché du travail. Cela permettrait d'établir clairement le lien entre leur mobilité et leurs parcours ainsi que leurs résultats sur le marché du travail, sans augmenter les coûts de la collecte ni le fardeau des répondants.
- c) Programme harmonisé de prêts aux étudiants (PHPE) :
Le couplage des données du PHPE à celles du SIAE et de l'END aiderait à comprendre les répercussions du programme de prêts sur la persévérance des étudiants, leurs études et leurs résultats sur le marché du travail. Il faut noter que l'administration du PHPE relève de Développement des ressources humaines Canada et que la base de données du PHPE réside à DRHC. Outre remplir les exigences de la politique sur le couplage des enregistrements de Statistique Canada, les couplages réalisés à des fins de statistiques et de recherche ne seraient faits que si DRHC consentait au couplage et à condition que Statistique extraie des données de la base de données du PHPE qui pourraient être utilisées à cette fin.
- e) Fichier d'impôt sur le revenu :
Pour le moment, les connaissances sur les gains à long terme des diplômés dans différentes disciplines sont incomplètes. Des enquêtes telles que l'Enquête sur la population active et le Recensement donnent de l'information approximative. Pour comprendre plus à fond cette question, le couplage du SISCOU à l'information fiscale fournie à Statistique Canada par l'Agence des douanes et du revenu du Canada (ADRC) en vertu des dispositions de la *Loi sur la statistique* a été approuvé dans le cadre de la politique sur le couplage des enregistrements de Statistique Canada. Même si elle donne des résultats très utiles, l'étude actuelle a ses limites. Le NAS est uniquement déclaré au niveau des universités dans les systèmes de données actuels et il est manquant pour des établissements, notamment tous ceux du Québec. Si le NAS était déclaré plus souvent dans le SIAE, nous pourrions avoir une image

complète et exacte des gains à long terme entre tous les niveaux d'enseignement postsecondaire des diplômés des différentes disciplines.

9 LA CONFIDENTIALITÉ, LA PROTECTION DE LA VIE PRIVÉE ET LA SÉCURITÉ DES RENSEIGNEMENTS

Toutes les enquêtes réalisées par le Centre de la statistique de l'éducation sur les effectifs et les diplômés postsecondaires recueillent des données sur le dossier de chaque étudiant. Or, contrairement à ses prédécesseurs, le SIAE demande des données signalétiques sur les étudiants, comme leur nom et leur adresse. Ces données soulèvent cependant des questions sur la protection de la vie privée.

Trois raisons motivent l'inclusion des données signalétiques sur les étudiants. Premièrement, c'est un moyen de communiquer avec les répondants aux enquêtes-échantillons. Deuxièmement, elles servent à établir un lien entre les données dans le temps. Troisièmement, elles assurent l'exactitude.

Mandat statistique et non administratif

Statistique Canada a un mandat statistique et non administratif. Contrairement aux bases de données ayant un rapport avec des fonctions administratives, comme l'immatriculation des véhicules automobiles qui touche directement les personnes visées, les données sur les étudiants conservées dans le SIAE, ainsi que les fichiers statistiques créés par couplage, ne visent aucune fin administrative.

Aucune décision n'est prise relativement à une personne à partir de l'information contenue dans les fichiers de données de Statistique Canada. Ces fichiers servent à produire des analyses de tendances et de profils. Une analyse a pour point de mire des groupements, plutôt que des personnes, et Statistique Canada garantit qu'aucune personne n'est identifiée dans les tableaux et les graphiques qui découlent d'un tel travail.

Coordonnées : Statistique Canada réalise des enquêtes de suivi auprès des étudiants et des diplômés. L'END est un exemple bien connu dans le milieu postsecondaire. Pour réaliser ces enquêtes, par le passé, Statistique Canada prélevait un échantillon à l'aide des identificateurs uniques attribués par les

collèges et les universités et inclus dans les enregistrements des étudiants. Statistique Canada demandait ensuite aux établissements des renseignements sur leur nom et leurs coordonnées.

Or, Statistique Canada et les établissements d'enseignement estiment que ce processus est long et coûteux. En vertu du SIAE, les renseignements sur son nom et les coordonnées nécessaires pour réaliser ces enquêtes seront dorénavant inclus dans l'enregistrement de chaque étudiant.

Couplage : L'un des objectifs fondamentaux du SIAE est de donner des renseignements sur la mobilité des étudiants entre les établissements ou les provinces pendant leurs études, qui peuvent s'étendre sur de nombreuses années. Contrairement aux enquêtes actuelles dont les enregistrements produits d'une année à l'autre ne sont pas couplés, l'enregistrement de chaque étudiant contenu dans le SIAE sera couplé de façon longitudinale afin de donner un historique cumulatif de la progression de cet étudiant dans le système d'enseignement postsecondaire. Cette mesure permettra d'étudier la mobilité, les parcours et les résultats de la manière décrite plus haut.

Exactitude : Le SIAE demande divers éléments signalétiques sur les étudiants pour assurer une exactitude élevée des données. La nécessité de l'exactitude est tout spécialement importante parce qu'il est fondamental que le SIAE donne des renseignements sur la mobilité des étudiants pendant plus de deux ou trois années consécutives.

L'inclusion des identificateurs personnels dans le SIAE a soulevé des questions quant au pouvoir de Statistique Canada d'obtenir des renseignements de même que sur la confidentialité, la protection de la vie privée et la sécurité des renseignements. La présente partie porte sur ces inquiétudes et des questions connexes, comme le couplage des enregistrements, l'avis aux étudiants de l'utilisation qui est faite des renseignements, l'objet du numéro national d'étudiant et celui du numéro d'assurance sociale, et la conservation des enregistrements.

Le pouvoir légal se rattachant au SIAE

En vertu de la *Loi sur la statistique*, Statistique Canada est tenu de réaliser un programme de statistiques sur l'éducation :

- L'article 3 confie à Statistique Canada le mandat de « recueillir, compiler, analyser, dépouiller et publier des renseignements statistiques sur les activités commerciales, industrielles, financières, sociales, économiques et générales de la population et sur l'état de celles-ci. »
- L'article 22 ordonne expressément à Statistique Canada de « recueillir, compiler, analyser, dépouiller et publier » des statistiques sur l'éducation.

- L'article 13 expose le pouvoir légal de Statistique Canada d'obtenir l'accès à des « documents ou archives conservés dans un ministère ou dans un bureau municipal, une personne morale, entreprise ou organisation », lorsque l'Agence exige ces archives aux fins de la *Loi*.

La confidentialité des renseignements

Pour équilibrer les pouvoirs de Statistique Canada de recueillir des renseignements, la *Loi sur la statistique* établit l'obligation juridique de protéger la confidentialité des renseignements sur les particuliers :

- L'alinéa 17(1)b) de la *Loi sur la statistique* interdit la communication des renseignements obtenus en vertu de la *Loi* « de telle manière qu'il soit possible, grâce à ces révélations, de rattacher à un particulier, à une entreprise ou à une organisation identifiable les détails obtenus dans un relevé qui les concerne exclusivement ».
- La *Loi sur la statistique* précise également que tous les employés de Statistique Canada doivent prêter le serment de ne pas divulguer de renseignements confidentiels liés à une personne, à une entreprise ou à une organisation. Une dérogation à ce serment entraînant une violation de la confidentialité constitue une **infraction criminelle** passible d'une amende et/ou d'une peine d'emprisonnement.
- L'article 24 de la *Loi sur l'accès à l'information* fédérale interdit au responsable d'une institution gouvernementale de divulguer un document demandé aux termes de la *Loi* s'il contient des renseignements dont la communication est restreinte en vertu d'une disposition d'une autre loi. L'article 17 est l'une de ces dispositions; les renseignements qui ne peuvent être communiqués aux termes de l'article 17 de la *Loi sur la statistique* ne peuvent pas non plus être divulgués en vertu de la *Loi sur l'accès à l'information*.

Tous les employés de Statistique Canada ayant accès à des renseignements statistiques confidentiels doivent se soumettre à une vérification de sécurité. Seuls les employés de Statistique Canada qui doivent prendre connaissance de fichiers confidentiels dans l'exercice de leurs fonctions sont autorisés à y avoir accès. En fait, l'accès à des données confidentielles se limite à un très petit nombre d'employés chargés de la maintenance des bases de données ou de travaux d'analyse et de recherche. S'il s'agit de fichiers couplés, seules les personnes qui travaillent au projet de recherche faisant appel aux fichiers couplés y ont accès.

La sécurité de la transmission des données du SIAE

Les données du SIAE seront transmises à Statistique Canada par Internet à l'aide de procédures de chiffrement protégées et approuvées par le gouvernement. Actuellement, les répondants utilisent un logiciel mis au point par Statistique Canada appelé *Outil de retour de données*. Ce logiciel chiffre les

enregistrements et les envoie à une adresse Internet spécifiée (de manière à ce qu'il ne puisse y avoir d'erreur de destination).

Statistique Canada met aussi à l'essai l'utilisation de l'ICP (infrastructure à clés publiques) pour la retransmission aux répondants des données de leur enregistrement afin qu'ils les vérifient et les corrigent. L'ICP est un protocole de transmission protégée largement utilisé dans le secteur bancaire et dans d'autres secteurs d'activité. Il fait appel au logiciel cryptographique très puissant Entrust. Il ne sert qu'à retourner des données confidentielles à leur fournisseur pour une vérification initiale et, plus tard, à remettre une copie des données finales du fournisseur une fois le traitement et la vérification terminés. Lorsqu'elles sont transmises à l'aide de l'ICP, les données ne peuvent être renvoyées qu'à un ordinateur bien précis et seule la personne autorisée qui en détient la « clé » peut y accéder. L'utilisation de l'ICP aide à réduire les coûts, tout en améliorant à la fois la qualité et l'opportunité des données.

La sécurité des renseignements

Tous les employés de Statistique Canada doivent assurer la sécurité des renseignements. L'Agence a mis en place un réseau de procédures et de systèmes de sécurité matérielle pour éviter que des personnes non autorisées n'aient accès à des renseignements confidentiels. Voici certaines de ses caractéristiques importantes :

- des systèmes d'entrée sélective sont établis et empêchent les personnes non autorisées d'avoir accès aux immeubles de Statistique Canada;
- des politiques interdisent au personnel de sortir des lieux de travail de Statistique Canada des renseignements statistiques confidentiels, en version papier ou électronique;
- les renseignements statistiques confidentiels sont conservés à Statistique Canada dans des endroits protégés;
- l'élimination des renseignements statistiques confidentiels s'effectue à Statistique Canada dans des conditions ne présentant aucun risque de fuite.

Pour assurer la protection des fonds de renseignements et de données, Statistique Canada a des réseaux interne et externe distincts, le réseau A et le réseau B. Même si les deux réseaux sont protégés au moyen de plusieurs fonctions intégrées dans le matériel et les logiciels, à titre de protection supplémentaire, la maintenance des données recueillies en vertu de la *Loi sur la statistique* n'est assurée que dans le réseau A. Ce réseau n'a pas de lien direct avec Internet, ce qui fait qu'il n'y a aucun accès extérieur indésiré ou imprévu. Une fois que les renseignements chiffrés du SIAE sont transmis du répondant au réseau B de Statistique Canada, ils sont transférés physiquement au réseau A avant d'être chiffrés, et aucun renseignement transféré ne reste dans le réseau B.

La confidentialité des produits statistiques

La principale base de données du SIAE est logée dans son propre serveur protégé situé dans une salle fermée à clé à laquelle seul un nombre limité d'employés de Statistique Canada a accès. Les analystes de Statistique Canada n'ont accès qu'à une base de données de recherche distincte qui est dépouillée des identificateurs personnels. Le logiciel de mise en tableaux fait également appel à des procédures automatisées d'arrondissement aléatoire et de suppression de cellules³. Toutefois, d'autres procédures sont en place :

- Tous les produits statistiques sont filtrés visuellement pour qu'il soit impossible d'y repérer des renseignements de nature délicate sur une personne. Cela signifie qu'un analyste à Statistique Canada examine manuellement tous les tableaux avant leur diffusion.
- Il existe aussi des procédures pour éviter l'identification d'un répondant par recoupements. L'identification d'un répondant par recoupements désigne le cas où deux fichiers de données distincts pourraient, s'ils étaient rassemblés, permettre d'identifier un particulier.

Ces procédures sont appliquées rigoureusement depuis des années dans le cas des bases de données collégiales et universitaires.

Le couplage des enregistrements

Il est reconnu que le couplage des enregistrements est un outil statistique valable et une solution de rechange aux données que doivent souvent soumettre les répondants. Néanmoins, parce que cette pratique peut soulever des questions sur la protection de la vie privée, depuis 1986, Statistique Canada a établi une politique sur le couplage des enregistrements. Cette politique vise à assurer le maintien d'un équilibre entre l'intérêt public, ce à quoi le couplage doit servir, et les risques d'atteinte à la vie privée. Les politiques et procédures de Statistique Canada garantissent le couplage des enregistrements si **toutes** les conditions suivantes sont respectées :

- l'objet de l'opération de couplage des enregistrements correspond au mandat de Statistique Canada tel qu'il est décrit dans la *Loi sur la statistique*;
- les produits de l'opération de couplage des enregistrements ne sont diffusés que conformément aux dispositions sur la confidentialité de la *Loi sur la statistique* et aux exigences applicables de la *Loi sur la protection des renseignements personnels*;
- le coût de l'opération de couplage des enregistrements ou les économies réalisées au niveau du fardeau du répondant par rapport à d'autres solutions de rechange peuvent être prouvés, ou cette opération est la seule option réalisable;

³ L'arrondissement aléatoire signifie que toutes les fréquences entre 1 et 9 sont arrondies à 0 ou 5 et ce, de manière aléatoire. La suppression des cellules signifie que toutes les cellules d'une fréquence inférieure à 5 (après l'arrondissement) ne peuvent être consultées par l'utilisateur.

- l'opération de couplage comporte des avantages qui sont nettement dans l'intérêt public;
- on juge que le couplage des enregistrements ne menace pas la réalisation ultérieure des programmes de Statistique Canada;
- le couplage respecte un processus prescrit d'examen et d'approbation établi et géré par Statistique Canada.

En outre, lorsqu'une opération de couplage d'enregistrements proposée engloberait le couplage d'enregistrements existants à de l'information que Statistique Canada recueillerait directement auprès des répondants, ces derniers doivent être avisés de l'opération de couplage proposée au moment de la collecte des données. Le couplage entre le SIAE et l'END en est un exemple. Ainsi, on pourrait demander aux répondants, dans le cadre des entrevues menées pour l'END, s'ils consentent à ce que l'enregistrement qui sera établi sur eux à la suite de l'enquête soit couplé à celui qui les vise dans le SIAE. L'objet du couplage et l'importance de l'information qui en découle leur seraient indiqués.

L'avis aux étudiants de l'utilisation qui est faite des renseignements les concernant

Conformément au principe du consentement éclairé pour assurer la protection de la vie privée, Statistique Canada collaborera avec les universités et les collèges afin d'élaborer une façon dont les établissements pourront aviser les étudiants de la manière dont les renseignements seront utilisés et de leurs droits relatifs à cet usage. Ils seront également avisés que leurs renseignements seront fournis à Statistique Canada en vertu des pouvoirs conférés par la *Loi sur la statistique*, à des fins de statistiques et de recherche. Il arrive parfois que les établissements le fassent par une mention dans leur annuaire. L'annexe A contient la formulation standard que Statistique Canada propose aux établissements.

Même si les enquêtes que le SIAE est en voie de remplacer permettraient également de recueillir des enregistrements sur les étudiants à l'aide d'identificateurs uniques, l'ajout d'information sur le nom et les coordonnées, combiné aux préoccupations de plus en plus grandes de la population en général au sujet de la protection de la vie privée et de la confidentialité des renseignements, a poussé Statistique Canada à demander aux établissements de donner aux étudiants plus d'information à ce sujet, de la manière décrite ci-dessus. Statistique Canada incite en particulier les établissements à inclure un renvoi au site Web du Bureau sur le SIAE, où ils peuvent en apprendre davantage sur le système et sur l'utilisation qui est faite de leurs données, et entrer en rapport avec Statistique Canada s'ils ont d'autres questions.

D'un point de vue juridique, lorsque des données sont tirées de dossiers administratifs en vertu de la *Loi sur la statistique*, les personnes que ces enregistrements visent n'ont pas le droit de refuser qu'ils soient diffusés. En

pratique, lorsqu'on le lui demandera, Statistique Canada supprimera de la base de données du SIAE les renseignements signalétiques visant un répondant (nom, adresse et NAS). Nous devons mettre en place des mécanismes pour garantir que les étudiants seront avisés de cette possibilité et nous tâcherons d'obtenir la collaboration des établissements pour qu'ils les fassent connaître. Une courte note sur les formules d'inscription des établissements et dans leur annuaire est une bonne façon de le faire; la formulation est proposée en annexe.

Le numéro national de l'étudiant du SIAE (NNE-SIAE)

Le NNE du SIAE est attribué de manière à permettre une identification *probabiliste* suffisante à des fins de recherche, plutôt qu'une identification *certaine* des répondants qui serait nécessaire pour utiliser ce numéro à des fins administratives. Le NNE-SIAE permet de coupler les enregistrements que les établissements envoient chaque année aux données sur les étudiants que la base de données nationale renferme déjà.

La capacité du SIAE à analyser la mobilité, les parcours des étudiants et leur lien avec leurs résultats dépend essentiellement du maintien du caractère unique des enregistrements sur les étudiants (il n'y a qu'un seul enregistrement pour chacun d'eux). Le couplage des enregistrements sur les étudiants à l'intérieur de la base nationale de données est la seule fin du NNE. Ce couplage ne vise aucune fin administrative. Contrairement à un numéro d'assurance-maladie, qu'il faut présenter pour obtenir des services de santé, les étudiants n'ont pas à utiliser leur numéro national pour obtenir des services d'enseignement. En réalité, il n'est pas nécessaire que les étudiants connaissent leur numéro national ni qu'ils s'en souviennent. Ce numéro est attribué par l'établissement que fréquente un étudiant suivant une formule établie par Statistique Canada et est reporté dans le système de dossiers des étudiants de l'établissement. Lorsqu'un établissement envoie des données à Statistique Canada, ce numéro est inclus dans l'enregistrement du SIAE. Quand l'étudiant s'inscrit dans un autre établissement d'enseignement postsecondaire, son numéro national est reporté sur son relevé de notes. L'établissement d'accueil reproduit le numéro national de l'étudiant dans son système de dossiers des étudiants à partir du relevé de notes de l'étudiant. Lorsque cet établissement envoie ses données à Statistique Canada, le NNE sert à établir le couplage à l'enregistrement existant pour l'étudiant. Statistique Canada disposera donc du seul fichier complet de NNE et le numéro national d'un étudiant ne sera utilisé dans aucun produit d'information du Bureau.

Le numéro d'assurance sociale (NAS)

Tout comme les enquêtes sur les effectifs qu'il remplacera, le SIAE comprend le numéro d'assurance sociale (NAS)⁴. L'une des caractéristiques principales du SIAE est sa nature longitudinale. Il faut une base de données longitudinale sur les étudiants pour comprendre leur mobilité et leur transition : dans le système

⁴ Environ 70 % des dossiers universitaires à l'extérieur du Québec renferment le NAS.

d'enseignement postsecondaire et à l'extérieur, entre les établissements et les provinces, et entre divers secteurs de l'enseignement postsecondaire. Le NAS sera utilisé de concert avec d'autres variables afin de coupler les enregistrements des étudiants dans le temps. Cela dit, même si les établissements et les organismes de coordination déclareront des données annuelles « instantanées » sur les étudiants, ces enregistrements seront couplés dans le temps dans la base de données du SIAE, au moyen du NAS et d'autres variables.

Pour étudier le mouvement des étudiants dans le système d'enseignement postsecondaire et à l'extérieur, d'un établissement à un autre et d'une province à une autre, les enregistrements doivent être couplés pour assurer un taux d'erreur très faible.

Le nom et les variables démographiques (comme la date de naissance et l'état civil) ne permettent pas une exactitude suffisante. Selon l'expérience de Statistique Canada, ces variables permettent d'établir des correspondances allant jusqu'à 80 % sur deux années consécutives. Toutefois, avec le temps, ce taux devient trop faible pour que les données soient utiles, surtout au cours des décennies qu'implique le concept de l'« acquisition continue du savoir ». Par conséquent, la nécessité du NAS a trait au principe de l'exactitude, c'est-à-dire que les renseignements personnels doivent être aussi exacts, complets et à jour qu'il le faut pour les fins auxquelles ils sont utilisés.

Le numéro d'assurance sociale permettrait également de faire un couplage exact entre un échantillon d'étudiants dans la base de données du SIAE et un échantillon de la base de données des prêts pour étudiants ou un échantillon des fichiers d'impôt. Par exemple, ces liens permettraient d'examiner la répartition du revenu par domaine d'études 10 ou 20 ans après l'obtention du diplôme. Tel que susmentionné, le NAS est nécessaire pour faire un couplage exact entre ces fichiers et sa présence comme identificateur commun dans les deux fichiers est capitale pour la stratégie prévue de couplage de ces fichiers dans un échantillon plutôt que sur une base de 100 %.

La conservation des enregistrements

Statistique Canada examine régulièrement la nécessité de conserver les enregistrements et des enregistrements sont détruits à intervalles réguliers. Dans le cas du SIAE, pour faire cet examen, il faut que les données du système aient trait à l'intérêt stratégique croissant manifesté à l'égard de l'acquisition continue du savoir. Les questions stratégiques supposent une certaine aptitude à suivre les études des particuliers pendant leur vie, même après de longs arrêts dans leur carrière d'étudiant.

L'accès des particuliers aux enregistrements qui les concernent

En vertu de la *Loi sur la protection des renseignements personnels* fédérale, les particuliers peuvent avoir accès à leurs propres renseignements qui sont conservés dans les banques de données fédérales, notamment ceux que détient Statistique Canada. Comme les bases de données existantes sur les étudiants ne renferment aucun nom, une personne ne peut avoir accès à ces enregistrements que si elle connaît le numéro d'étudiant que lui a attribué l'établissement qu'elle fréquente. Le SIAE renfermera le nom de la personne et pourra être consulté directement à l'aide de cette variable.

L'accès aux données du SIAE aux fins d'analyse

Accès à Statistique Canada : La principale base de données du SIAE renferme des renseignements sur le nom et les coordonnées ainsi que des identificateurs uniques. L'accès à cette base de données se limite aux employés de Statistique Canada qui ont réussi une vérification de fiabilité approfondie et qui ont obtenu l'autorisation spécifique de consulter directement les données. Ce petit nombre d'employés est chargé de garder les données et de veiller à ce que les renseignements consultés dans la base de données ne permettent pas d'identifier les particuliers. Comme nous l'avons déjà mentionné, par mesure de sécurité, entre autres choses, la base de données se trouve dans un serveur protégé situé dans une pièce fermée à clé.

Les chercheurs de Statistique Canada n'ont pas directement accès à la principale base de données du SIAE. Ils font des recherches à partir des bases de données de recherche qui ne contiennent pas de renseignements d'identification personnelle.

Pour que personne ne puisse être identifié dans les produits statistiques, le logiciel utilisé pour interroger la base appelle automatiquement l'arrondissement aléatoire et la suppression de cellules. De plus, des analystes spécialisés effectuent une vérification visuelle des produits. En pratique, cela signifie que chaque tableau est examiné et qu'une décision d'expert est prise pour veiller à ce qu'aucun particulier ne puisse être identifié.

Accès des fournisseurs de données : Les fournisseurs de données (les établissements et, dans certains cas, les gouvernements ou les organismes) recevront les données sur leurs étudiants ainsi que d'autres variables dérivées qui leur seront retournées. De cette manière, les établissements disposeront d'une copie de l'ensemble complet des données que Statistique Canada conservera sur eux. À l'aide de ces données et du logiciel que Statistique Canada procurera à chaque établissement, ce dernier sera en mesure d'effectuer la même analyse sur ses propres données que l'Agence est en mesure de faire sur les données nationales. Les fournisseurs de données pourront ainsi utiliser leurs propres données et le logiciel afin d'élaborer des interrogations pour comparer leurs propres données aux résultats non

confidentiels d'autres fournisseurs de données. Ils pourront ensuite envoyer ces interrogations par courriel à Statistique Canada qui les exécutera dans sa base de données de recherche nationale. Les données seront vérifiées visuellement par souci de confidentialité et retournées au demandeur.

Accès par d'autres : Les gouvernements, les organismes, les chercheurs indépendants et le public qui n'apportent pas de données mais qui ont besoin d'avoir accès aux données recevront une base de données fictives⁵ et le logiciel leur permettant d'élaborer leurs interrogations. De même, ils pourront envoyer leurs interrogations par courriel à Statistique Canada où elles seront exécutées dans la base de recherche et filtrées à des fins de confidentialité avant d'être retournées au demandeur.

La conception d'un fichier de microdonnées à grande diffusion sera aussi étudiée. Les « fichiers de microdonnées à grande diffusion » sont des fichiers d'enregistrements individuels dépouillés de tout identificateur et examinés par un comité technique interne de Statistique Canada pour garantir que les enregistrements ne contiennent aucune variable dans un format qui permettrait d'identifier à nouveau le répondant.

Un rapport connexe, *Système d'information amélioré sur les étudiants : réponses aux inquiétudes relatives à la protection de la vie privée et des renseignements personnels*, présente en détail l'engagement de Statistique Canada envers les normes nationales qui existent actuellement à l'égard de la protection de la vie privée.

10 PRODUITS ET SERVICES

Statistique Canada publiera les données du SIAE sous forme de tableaux statistiques normalisés et de documents de recherche. Les tableaux normalisés paraîtront dans Éducation au Canada et pourront être obtenus toute l'année auprès du Centre de la statistique de l'éducation. Les documents de recherche seront publiés dans la Revue trimestrielle de l'éducation. Les données du SIAE serviront également à produire plusieurs des indicateurs qui paraissent maintenant dans Indicateurs de l'éducation au Canada ainsi que dans la publication annuelle de l'OCDE intitulée Regards sur l'éducation. Les publications de Statistique Canada sont disponibles sur abonnement et le public peut aussi les consulter sans frais grâce au programme fédéral des bibliothèques de dépôt.

⁵ Une base de données « fictive » contient des données « simulées » qui ressemblent aux vraies données. Comme les bases de données « fictives » et « réelles » ont la même structure, un client peut faire des interrogations dans la base de données qui sont exécutées dans la base de données de recherche « réelle » par Statistique Canada.

Les travaux réalisés pour améliorer l'accessibilité des données aux établissements et aux autres répondants (comme les ministères de l'éducation) sont également très avancés. Ces groupes recevront leurs propres données en plus des variables dérivées et d'un outil d'interrogation mis au point par Statistique Canada.

À l'aide de cet outil et à partir de leurs propres données, ils pourront formuler des interrogations. Les interrogations seront envoyées par courriel à Statistique Canada, exécutées dans la base de recherche du SIAE puis retournées à leur expéditeur. La base de recherche sera dépourvue d'identificateurs personnels et programmée de manière à lancer l'arrondissement aléatoire et la suppression de cellules pour assurer la confidentialité. Les données seront également vérifiées visuellement à des fins de confidentialité.

Les utilisateurs du SIAE qui ne fournissent pas de données obtiendront une base de données fictive et un outil d'interrogation. Ils seront donc en mesure de formuler leurs propres interrogations qui seront appliquées à la base de recherche à Statistique Canada.

Les chercheurs universitaires et divers organismes qui ont besoin de données pour élaborer des politiques, des organismes comme l'AUCC et l'ACCC, pourraient également communiquer avec Statistique Canada pour obtenir des analyses et des produits de données personnalisés. Ces demandes seront traitées au cas par cas et les services seront offerts selon le principe du recouvrement des coûts.

Par ailleurs, la création d'un fichier de microdonnées à grande diffusion sera étudiée.

Tel que décrit dans la partie 2, Statistique Canada, sous réserve des dispositions prévues dans la *Loi sur la statistique*, échangera les données sur les établissements et les programmes avec plusieurs autres organismes. L'Agence songera aussi à élaborer une base de données accessible au public contenant ces renseignements.

Pour obtenir de plus amples renseignements sur le SIAE, consultez les adresses suivantes :

http://www.statcan.ca/francais/concepts/ESIS/index_f.htm

<http://www.statcan.ca/english/concepts/ESIS/index.htm>

ANNEXE

ÉBAUCHE DE TEXTE À INCLURE DANS LES FORMULES D'INSCRIPTION ET LES ANNUAIRES

FORMULES D'INSCRIPTION

En vertu de la *Loi sur la protection des renseignements personnels* fédérale, les particuliers peuvent demander l'accès à leurs propres renseignements qui sont détenus dans des banques d'information fédérales, notamment celles de Statistique Canada.

Les étudiants qui ne désirent pas que les renseignements les concernant soient utilisés peuvent demander à Statistique Canada de retirer de la base de données leurs renseignements signalétiques.

ANNUAIRES

Avis de divulgation de renseignements personnels à Statistique Canada

Statistique Canada est le bureau de la statistique nationale. À ce titre, Statistique Canada réalise des centaines d'enquêtes chaque année sur une vaste gamme de questions, notamment l'éducation.

Il est essentiel de pouvoir suivre les étudiants dans le temps et les établissements pour comprendre, par exemple, les facteurs qui influent sur la demande d'effectifs dans les établissements postsecondaires. Comme la reddition de comptes attire de plus en plus l'attention quand on parle des investissements publics, il est également important de comprendre les « résultats ». Pour effectuer ce genre d'études, Statistique Canada demande à tous les collèges et universités de lui fournir des données sur les étudiants et les diplômés. Les établissements recueillent des renseignements signalétiques sur les étudiants (nom, numéro d'identification de l'étudiant, numéro d'assurance sociale) et sur ses coordonnées (adresse et numéro de téléphone), des caractéristiques démographiques, des renseignements sur les effectifs, les études antérieures et la participation au marché du travail, avant de les remettre à Statistique Canada.

La *Loi sur la statistique* fédérale confère à Statistique Canada le pouvoir juridique nécessaire pour obtenir l'accès aux renseignements personnels que détiennent les établissements d'enseignement. Ces renseignements peuvent être utilisés uniquement à des fins statistiques et les dispositions sur la confidentialité figurant dans la *Loi sur la statistique* empêchent la communication des renseignements de quelque façon que ce soit qui permettrait d'identifier un étudiant.

Les étudiants qui ne désirent pas que les renseignements les concernant soient utilisés peuvent demander à Statistique Canada de retirer de la base de données leurs coordonnées personnelles.

Pour en savoir plus sur l'utilisation de ces renseignements, vous pouvez consulter le site Web de Statistique Canada à l'adresse www.statcan.ca/francais/concepts/ESIS/index_f.htm, écrire à l'adresse suivante: Section de l'éducation postsecondaire, Centre de la statistique de l'éducation, 17e étage, Immeuble R.H. Coats, ParcTunney, Ottawa, K1A 0T6 ou téléphoner au 613-951-1666.