

Vers un système statistique centré sur des registres : Éléments nouveaux à Statistique Canada

Jean Pignal, Philippe Gagné, Christian Wolfe et Tanvir Quadir¹

Résumé

Le Projet de transformation et d'intégration des registres statistiques (PTIRS) vise à bâtir et à entretenir une infrastructure de registres statistiques (IRS) se composant de quatre registres de base (population, immeubles, entreprises et activités) entièrement créés par l'intégration des données administratives de diverses sources qui seront utilisées à des fins statistiques. L'IRS a pour but de maximiser l'utilisation des renseignements utiles tirés des données déjà recueillies, d'améliorer l'actualité des statistiques et de réduire le fardeau de réponse tout en protégeant la vie privée des Canadiens. Une brève description de l'élaboration en cours d'un registre de la population anonymisé et d'un registre des immeubles, du remaniement du registre existant des entreprises, et du cadre conceptuel d'un registre des activités sera fournie. Le présent article décrit également le Système canadien de registres statistiques intégrés (SCRSI), lequel assemble les registres de base sous-jacents et pose le fondement de l'IRS.

Mots clés : Registre statistique, données administratives, anonymisation, infrastructure de registres statistiques

1. Introduction

Comme on le voit dans un certain nombre de pays développés, l'utilisation accrue de données administratives est une composante centrale de nombreuses stratégies de modernisation pour Statistique Canada². Dans cette optique, le Projet de transformation et d'intégration des registres statistiques (PTIRS) vise à maximiser l'information des données administratives qui ont été réutilisées à des fins statistiques. Cette pratique réduit ensuite le fardeau de réponse et le coût de la collecte de données d'enquête et permet aux utilisateurs d'accéder à des renseignements statistiques opportuns, pertinents et de qualité dans un fonds de données anonymisé, intégré et central. Ainsi, Statistique Canada procède à la modernisation de l'infrastructure qui interconnectera une série de registres statistiques de base (population, immeubles, entreprises et activités), laquelle sera utilisée comme plateforme cohérente par tous les programmes statistiques de l'organisme. Il s'agit de l'élaboration d'un Registre statistique de la population (RSP), de la transformation du Registre des adresses (RA) résidentielles existant en un Registre statistique des immeubles (RSI) plus complet et de l'ajout de nouvelles composantes au Registre statistique des entreprises (RSE). En outre, dans le cadre des initiatives de modernisation des statistiques sociales, la version de validation de principe d'un Registre statistique des activités (RSA), le fichier analytique de la durée de vie (LAF), est en cours d'élaboration.

¹Jean Pignal, Division de l'infrastructure d'intégration des données, Statistique Canada, Canada, K1A 0T6 (jean.pignal@canada.ca); Philippe Gagné, Division de l'infrastructure d'intégration des données, Statistique Canada, Canada, K1A 0T6 (philippe.gagne5@canada.ca); Christian Wolfe, Division de l'infrastructure d'intégration des données, Statistique Canada, Canada, K1A 0T6 (christian.wolfe@canada.ca); Tanvir Quadir, Division des enquêtes spéciales, Statistique Canada, Canada, K1A 0T6 (tanvir.quadir@canada.ca). ²D'autres pays se penchent actuellement sur l'utilisation accrue des données administratives à des fins statistiques : l'Australie, la Nouvelle-Zélande, le Royaume-Uni et les États-Unis.

²D'autres pays se penchent actuellement sur l'utilisation accrue des données administratives à des fins statistiques : l'Australie, la Nouvelle-Zélande, le Royaume-Uni et les États-Unis.

2. Ensemble de registres statistiques de base

2.1 Registre statistique des entreprises

Le RSE de Statistique Canada a d'abord été créé dans les années 1980, et a fait l'objet d'améliorations continues au fil du temps. Conformément aux lignes directrices internationales³, Le RSE recueille, compile et entretient un registre complet des entreprises (entités juridiques et statistiques) au Canada afin de fournir une base complète, simple et à jour pour les programmes économiques. Il joue également un rôle central dans la modernisation de l'approche de l'organisme à la production de statistiques économiques.

Le RSE a d'abord été élaboré comme base pour l'échantillonnage d'enquêtes et est utilisé à l'heure actuelle par plus de 200 enquêtes économiques. Son principal rôle est de soutenir les activités de collecte de données d'enquête par la gestion et la mise à jour des renseignements sur les entreprises et les répondants, et la surveillance et le contrôle du fardeau de réponse des répondants. La couverture cible du RSE inclut toutes les entreprises et institutions, soit plus de 6,6 millions d'entités actives, qui produisent des biens ou des services au Canada, y compris les entreprises agricoles et les fermes.

À l'instar du RSI et du RSP proposés, le RSE est évolutif, car il est continuellement mis à jour au moyen de différentes sources d'information et repose principalement sur les données administratives. Une très grande portion du RSE est entretenu uniquement par l'utilisation de fichiers de données administratives hebdomadaires, mensuels, trimestriels et annuels de l'Agence du revenu du Canada. Les activités d'établissement du profil des entreprises sont essentielles à la création de structures d'entreprise complètes. Enfin, les commentaires reçus pendant les enquêtes complètent les données administratives et les activités d'établissement de profils.

Pour ce qui est des attributs, le RSE contient des identificateurs directs tels que les dénominations sociales et les noms commerciaux des unités économiques. Toutes les entités dans le RSE se voient attribuer un identificateur statistique unique d'entreprise (BR_SN) afin d'éviter les redondances et de permettre l'anonymisation des entités. Le registre contient également des variables statistiques comme la date de « naissance » (date de constitution en société ou de la reconnaissance officielle comme exploitant économique) et de cessation, l'état des activités de l'unité économique, les variables relatives à la taille (par exemple, les ventes ou le nombre d'employés) et les variables relatives à la classification comme le SCIAN et le CCUSI⁴, entre autres. Toutes les valeurs comportent des références temporelles. Les variables relatives à la communication sont également présentes, y compris la personne-ressource, l'adresse postale, les numéros de ligne téléphonique terrestre et de cellulaire, ainsi que les adresses de courriel et des sites Web des entités.

Grâce à l'accessibilité et à l'utilisation accrues des fichiers de données administratives, le RSE est en voie de devenir un carrefour central pour le couplage de données d'entreprise, une entrée directe pour l'analyse des populations des entreprises et pour la diffusion de statistiques officielles sur les entreprises. Le RSE optimise l'élaboration du système statistique centré sur des registres en ajoutant la relation aux unités présentes dans les autres registres statistiques de base – par exemple, l'emplacement des entités du RSE sera bientôt enregistré au moyen des identificateurs statistiques d'immeubles et d'unités de bâtiment dérivés du RSI. En outre, afin de mieux répondre aux besoins analytiques, le RSE est en train de créer un nouveau module longitudinal. Il inclura des identificateurs longitudinaux, des tableaux de prédécesseurs et de successeurs comportant des renseignements détaillés, ainsi que des mises à jour historiques sur les données antérieures afin de tenir compte des renseignements accessibles les plus récents. Les nouvelles caractéristiques du RSE doté de couplages avec les autres registres statistiques de base fourniront à l'infrastructure une réponse rapide aux besoins changeants en matière de données longitudinales détaillées sur les entreprises et faciliteront la production de statistiques sur les entreprises et les indicateurs d'entrepreneuriat, ainsi que des statistiques nouvelles et innovantes.

³ UNECE. (2015). Guidelines on Statistical Business Registers. Récupéré sur https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf

⁴ Le SCIAN est le Système de classification des industries de l'Amérique du Nord et le CCUSI est la Classification canadienne des unités et secteurs institutionnels.

2.2 Registre statistique des immeubles

Le Registre des adresses (RA), créé en 1986, a été utilisé comme fondement des collectes du recensement et comme base pour les enquêtes sur les ménages et les enquêtes sociales. Or, le rôle de ce registre semble ne pas être aussi crucial que celui du RSE. Par exemple, le RA ne consigne pas les unités échantillonnées pour le monitoring du fardeau des répondants et n'a pas ou peu d'interaction directe en ce qui a trait aux activités de collecte et du processus d'enquêtes. Ainsi, dans le cadre des initiatives de transformation, le RA sera transformé en un registre statistique des immeubles (RSI) plus complet qui couvrira les immeubles résidentiels et non résidentiels, où l'entité est le bâtiment et ses unités. Le terme « Bâtiment » renvoie à une structure permanente autostable couverte d'un toit, généralement fermée par des murs extérieurs ou des murs de séparation qui vont des fondations jusqu'au toit, et qui comporte une ou plusieurs pièces, ou d'autres espaces, alors que le terme « unité de bâtiment » renvoie à un lieu qui fait partie d'un bâtiment et est soit résidentiel ou non résidentiel. L'unité de bâtiment doit avoir sa propre entrée soit par une porte extérieure ou par une porte intérieure dans un couloir partagé. Un numéro statistique unique et stable sera attribué à chaque immeuble (Bg_SN) et unité de bâtiment (BU_SN) et sera utilisé pour coupler aux registres de base tout autre registre statistique ou fichier administratif de l'organisme et pour protéger les renseignements personnels.

Le RSI sera mis à jour de façon évolutive et interreliée avec le RSP et le RSE. Le RSI sera soutenu par l'Infrastructure statistique géospatiale (ISG)⁵ pour améliorer le géocodage des fichiers administratifs et optimiser le couplage des données de géolocalisation administratives (p. ex. les adresses) avec le RSI et indirectement avec le RSP et le RSE. Certains des fichiers administratifs faisant actuellement l'objet d'analyses aux fins d'inclusion dans le RSI sont les fichiers de la Société canadienne des postes (déjà utilisés dans le RA), les fichiers provinciaux des services d'urgence, les fichiers d'évaluations foncières et les fichiers de registres fonciers, ainsi que les fichiers sur les compteurs intelligents provenant des entreprises hydroélectriques provinciales.

Les attributs du RSI sont classés dans les quatre catégories ci-après :

- Géographie (îlot de base, côté d'îlot, coordonnées GPS, secteurs géographiques de diffusion et de collecte)
- Coordonnées (adresse municipale, adresse postale, numéro de téléphone) à utiliser pour les activités de collecte et de couplage
- État (inactif, actif, temporaire, substitut)
- Indicateurs de qualité (validité, occupation résidentielle, précision géographique, etc.)

2.3 Registre statistique de la population

Le Registre statistique de la population (RSP) est une base de données évolutive, simple et hautement protégée de toutes les personnes se trouvant dans divers fichiers administratifs à compter de 2016. Même si les fichiers administratifs sont utilisés depuis longtemps au sein de l'organisme, l'idée d'un registre statistique de la population longitudinal provient de l'utilisation des données administratives dans le programme de Recensement de la population afin de remplacer la collecte des données sur le revenu. Le fait que la couverture de la population canadienne provenant de l'intégration des fichiers administratifs existants s'est avérée très vaste, a encouragé l'équipe du recensement à entreprendre un projet de recherche, la base statistique de données démographiques canadiennes (BSDDC), afin de mesurer notre capacité à reproduire le dénombrement des Canadiens en plaçant les bonnes personnes au bon endroit, uniquement au moyen des données administratives. En tout, cinq versions de la BSDDC ont été créées depuis 2013, dont deux représentent le Recensement de 2011 et trois représentent le Recensement de 2016. La dernière version de cette base de données sera utilisée pour amorcer le RSP avec de nouveaux fichiers entrants qui mettent continuellement à jour le registre en ajoutant de nouveaux enregistrements et en ajoutant de nouvelles valeurs d'attribut aux enregistrements existants. Les fichiers administratifs ayant contribué à la BSDDC (et au futur RSP) sont divers fichiers fiscaux des particuliers obtenus auprès de l'Agence du revenu du Canada (ARC), les naissances et les décès provenant des statistiques de l'état civil, les fichiers des résidents temporaires et permanents d'Immigration, Réfugiés et

⁵ L'ISG est un sous-ensemble de l'Infrastructure des données spatiales (IDS) pour ces concepts clés : îlots de base (IB), rues, côté d'îlot (CI) et subdivision de recensement (SDR). D'autres précisions au sujet de l'IDS sont indiquées dans la section des références.

Citoyenneté Canada (IRCC), le Registre des Indiens de Relations Couronne-Autochtones et Affaires du Nord Canada (RCAANC), et le registre des numéros d'assurance sociale (NAS) d'Emploi et Développement social Canada (EDSC). Outre les fichiers de la BSDDC susmentionnés, on prévoit que, dès que le RSP sera parvenu à maturité, tous les nouveaux fichiers administratifs portant sur des personnes pourraient être filtrés par le RSP. Certains de ces fichiers, comme les fichiers des permis de conduire et des services de téléphonie cellulaire, pourraient ajouter des enregistrements et des attributs au RSP, alors que d'autres pourraient simplement devenir anonymisés et se voir attribuer un numéro statistique individuel (NSI).

La raison d'être du NSI est de remédier aux préoccupations relatives à la protection des renseignements personnels en retirant les identificateurs personnels (noms, numéros administratifs comme le NAS, etc.) de tous les nouveaux fichiers administratifs et de leur substituer un numéro aléatoire permanent qui permettrait à l'organisme de continuer à utiliser l'information à des fins statistiques en permettant les couplages d'enregistrements tout en masquant l'identité des personnes liées aux fichiers. Le but ultime est de préserver la confiance du public, alors que Statistique Canada acquiert des sources de données administratives susceptibles de créer des statistiques plus opportunes et pertinentes.

Les attributs démographiques inclus dans le RSP sont très élémentaires : sexe, statut d'immigrant, statut d'Autochtone, date de naissance, date de décès, revenu (à des fins de stratification), état civil, date d'immigration et d'émigration. Les coordonnées incluent le nom, l'adresse (remplacé par le BU_SN provenant du RSI), le numéro de cellulaire, l'adresse de courriel et la langue de correspondance (pour les possibles communications futures). Ces attributs sont principalement divisés en attributs statiques, comme la date de naissance, qui ne changent pas habituellement, et en attributs dynamiques, comme le nom de famille, qui peuvent varier dans le temps. Le RSP entretient des liasses de valeurs pour chaque attribut provenant de sources disparates et a élaboré des règles opérationnelles pour les valeurs les plus probables de chaque attribut devant être utilisé par les programmes statistiques. Pour ce faire, le RSP a mis en œuvre les indicateurs de la qualité suivants :

- État du NSI/validité de la personne : Il s'agit d'un indicateur du degré de confiance à l'idée qu'une personne est réellement une entité statistique unique dans le RSP. Le NSI peut être désigné comme temporaire (« T »), permanent (« P ») ou échu (« X ») en fonction des signaux reçus des fichiers administratifs sous-jacents.
- Lien valide/indicateur de validité : Cet indicateur de validité est une valeur binaire (0,1) ou (V, F) au niveau de l'identificateur. Cela sera utilisé pour gérer les faux positifs lorsque l'information est attribuée par erreur au mauvais enregistrement (NSI).
- Qualité d'attribut : Valeur calculée ou attribuée en fonction de la confiance que nous avons en la source pour un attribut précis.
- Indicateur dans le champ ou hors du champ de l'enquête : Cette valeur est calculée au moment de l'extraction pour déterminer si, compte tenu des paramètres d'extraction (période, groupe d'âge souhaité, proximité du dernier signal enregistré, indication d'émigration, date de décès, etc.), l'enregistrement est dans le champ ou hors du champ de l'enquête.
- Entropie : Il s'agit d'une mesure de la divergence et de la disparité dans toutes les sources enregistrées relativement aux attributs d'un enregistrement donné. Plus l'entropie s'approche d'une valeur nulle, plus l'on considère qu'il y a correspondance entre les sources quant à la valeur de l'attribut. La valeur pondérée de l'entropie est calculée de manière dynamique lorsqu'on exécute une extraction du RSP. En ce qui a trait aux attributs dynamiques, le nombre d'occurrences d'une valeur ainsi que sa proximité temporelle sont également pris en considération.

2.4 Registre statistique des activités (RSA)

L'idée d'un Registre statistique des activités (RSA) figure dans le modèle de la Commission économique des Nations Unies pour l'Europe (CEE-ONU) (2013)⁶ pour les Nordic Register-based Statistical Systems (systèmes statistiques

⁶ CEE-ONU (2013) « Using Administrative Data in Statistical Registers » (utilisation des données administratives dans les registres statistiques). Récupéré sur : <https://statswiki.unece.org/display/adso/Using+Administrative+Data+in+Statistical+Registers> (en anglais).

nordiques fondés sur des registres) et dans Wallgren et Wallgren (2014). L'idée d'un registre statistique intitulé « Social Data Register » (registre des données sociales [SDR]) considéré comme contributeur potentiel au Registre statistique des activités est d'abord née du plan stratégique de l'organisme en vue des initiatives de modernisation du programme de la statistique sociale (MacNabb, 2017). La version pilote du SDR, intitulée « Life-course Analytical File » (fichier analytique de la durée de vie [LAF]), est en cours, laquelle consiste en un registre sur les activités individuelles pour la population canadienne dans les domaines suivants : travail, études et capital humain, bien-être économique, famille, crime et victimisation, santé et bien-être, immigration et statut autochtone. Cooke (2011) a brossé un tableau de l'intégration selon le point de vue de la durée de vie à l'intérieur du système canadien de données sociales dans un article préparé pour Emploi et Développement social Canada (EDSC) et Statistique Canada. Cette référence fondamentale jette les bases du LAF qui serviront à effectuer ce qui suit :

- contribuer éventuellement à la création d'un registre statistique des activités, lequel formera le quatrième registre statistique de base dans le Système canadien de registres statistiques intégrés (SCRSI) après les registres statistiques de la population, des immeubles et des entreprises;
- faciliter l'intégration des sources sous-jacentes avec diligence raisonnable en ce qui concerne la sécurité et la confidentialité des données en créant un registre statistique anonymisé contenant des données dans les domaines mentionnés ci-dessus, lesquelles ne sont pas préservées dans d'autres registres de base (population, immeubles et entreprises);
- encourager la recherche statistique interdisciplinaire en fournissant un fichier prêt à la réalisation d'analyses, qui comprend les variables et les indicateurs transdisciplinaires clés requis par la plupart des programmes de statistique sociale;
- enregistrer et maintenir les activités historiques de la population canadienne qui permettront d'analyser le point de vue de la durée de vie, les trajectoires développementales et l'histoire des événements;
- contribuer au soutien amélioré en matière d'enquête (service d'encadrement, remplacement de contenu);
- la couverture du LAF s'étend de 2005 à 2015. Le LAF recourt à l'Environnement de couplage de données sociales (ECDS) pour coupler des enregistrements d'un domaine à l'autre et consigner les activités pertinentes. La liste de sources administratives utilisées pour créer le LAF inclut, par domaine et par sujet : le travail – fichiers fiscaux et fichier de la prestation d'assurance-emploi; études et capital humain – enregistrements des établissements d'enseignement postsecondaire et d'apprentissage; bien-être économique – fichier d'impôt (rendement économique détaillé); famille – fichier d'impôt (données sur les familles); crime et victimisation – données des tribunaux de juridiction criminelle; santé et bien-être – données sur l'état civil, données du Registre des cancers et d'hospitalisation; immigration et statut autochtone – fichiers sur les résidents permanents et non permanents et le Registre des Indiens.

La couverture du LAF s'étend de 2005 à 2015. Le LAF recourt à l'Environnement de couplage de données sociales (ECDS) pour attribuer de manière sécurisée les activités à des personnes anonymisées d'un domaine à l'autre. Les domaines seront traités séparément et ne seront pas combinés dans une base de données. Une grande catégorie d'attributs du LAF est fournie ci-dessous, dans laquelle certains des attributs peuvent se chevaucher avec ceux du RSP, lesquels ne seront pas maintenus par le LAF à long terme afin d'éviter les incohérences.

- Travail : situation d'emploi, revenu d'emploi, prestations d'assurance-emploi
- Études et capital humain : inscription, type et durée du programme ou du certificat, établissement d'enseignement, niveau de scolarité
- Famille : Composition de la famille (recensement), revenu de famille, type de ménage, revenu de ménage (si l'indicateur de ménage est accessible à partir du Registre des adresses)
- Santé et bien-être : Hospitalisation, utilisation de services de santé, maladie chronique, données sur les naissances et les décès
- Bien-être économique : Rendement économique détaillé (revenu d'emploi – salaires, revenu d'un travail indépendant, revenu de placements, revenu privé de retraite, autre revenu du marché, transferts gouvernementaux)
- Crime et victimisation : Types et date des infractions et des accusations, incarcération, récidivisme
- Immigration et statut autochtone : Caractéristiques des immigrants, indicateur autochtone

3. Infrastructure de registres statistiques

Le mandat de l'infrastructure de registres statistiques (IRS) est de créer un cadre pour les registres statistiques de base intégrés qui sera utilisé comme assise cohérente pour les programmes statistiques et pour exécuter les processus de couplage des registres à l'appui du rapprochement de la base, de l'intégration entre registres, du géocodage et de l'anonymisation des fichiers administratifs.

Une division administrative de Statistique Canada est chargée de l'acquisition et de la réception du fichier de données administratives (FDA), de son traitement préalable, de déterminer les champs comportant des identificateurs et des quasi-identificateurs, de préparer les métadonnées, de stocker le FDA et de contrôler l'accès à ce dernier. Les FDA traités dans la division administrative subiront le géocodage spatial si le fichier de données contient des attributs géographiques non tabulaires ou est un fichier spatial. La phase de transformation préalable des données inclut tout traitement supplémentaire nécessaire pour transformer les données dans les formats attendus par le moteur d'appariement aux registres (MAR). L'intégration d'un FDA consiste à définir la disposition des renseignements dans le FDA, à associer les variables dans le FDA pour en enregistrer les champs et à effectuer l'évaluation de la qualité. Le MAR fournit à l'IRS les outils lui permettant d'exécuter le couplage de données vers le système de registres statistiques ou à l'intérieur de ce dernier. Le MAR comportera les éléments fondamentaux ci-après :

- attribuer un numéro statistique unique à chaque nouvelle unité statistique (personne, immeuble, unité de bâtiment, entreprise ou activité);
- dépersonnaliser les unités en cachant les identifiants personnels bruts retirés du registre et en les stockant dans un environnement sécurisé et contrôlé
- élaborer des règles d'affaires afin de déterminer avec une certitude raisonnable si un nouvel enregistrement des sources administratives implique véritablement une nouvelle unité;
- placer les personnes dans leur logement. Il y aura des demandes d'information du RSP au RSI pour obtenir l'identificateur unique. Cela requiert une solide base de données géo-spatiale qui contient les informations relatives au réseau routier, aux tranches d'adresses et à toutes les frontières géographiques prises en comptes à Statistique Canada.

4. Avantages et utilisations du système de registres statistiques

Alors que la conception et la construction du nouveau Registre statistique de la population et du Registre statistique des immeubles sont en cours d'élaboration d'abord pour appuyer et compléter les efforts actuels de collecte du recensement, il existe d'autres utilisations statistiques potentielles pour de tels registres au sein de l'organisme de statistique national. Par exemple, le RSP pourrait être utilisé pour compléter notre actuel Programme des estimations démographiques. De plus, lorsque la qualité est suffisamment élevée, il pourrait fournir, pour la première fois, une base de sondage des personnes pour les programmes de statistique sociale nous permettant de bâtir des plans d'échantillonnage plus efficaces. Cela compléterait la base de logements fournie par le RSI. En outre, cela produirait des bases de référence régulières pour les études de la population (p. ex. instantanés de la population entière à un moment précis, ou d'un sous-ensemble de la population [jeunes, Canadiens autochtones, nouveaux immigrants] ou pour une géographie précise).

Le terme « intégration » dans le SCRSI devrait être perçu sur le plan de l'interactivité des registres statistiques de base entre eux et avec les registres statistiques non de base, alors que les registres statistiques de base contiennent des renseignements très élémentaires au sujet des unités statistiques et que les registres satellites ou non de base contiennent un certain nombre de précisions sur la construction ou les concepts sous-jacents d'intérêt (p. ex. l'emploi, la santé, la justice, l'éducation).

5. Conclusion et orientations futures

La raison d'être du système statistique centré sur les registres repose sur le modèle « les données administratives d'abord » de la stratégie de modernisation de Statistique Canada. Le Recensement de la population examine

actuellement le moyen de remplacer l'ensemble ou certaines parties de la collecte traditionnelle de données au moyen des données administratives. La poursuite de la simulation du recensement combiné (c'est-à-dire l'IRS combiné au dénombrement traditionnel) jusqu'en 2021 donnera lieu à une décision relative au format du recensement à partir de 2026. Dès que les registres statistiques de base seront fonctionnels, il sera essentiel d'assurer leur intégration sans heurts aux registres statistiques non de base (satellites). Il faut élaborer les lignes directrices et les cadres pour produire des statistiques officielles au moyen du système statistique centré sur les registres ou un mélange de registres statistiques et d'enquêtes ainsi que des indicateurs de la qualité pour aider à interpréter les résultats.

Bibliographie

- Bakker, B. F. M. et al. (2014), "The System of Social Statistical Datasets of Statistics Netherlands: An Integral Approach to the Production of Register-based Social Statistics", *Statistical Journal of the IAOS*, 30, pp. 411-424.
- Cooke, M. (2011), "Integrating a Lifecourse Perspective within the Canadian Social Data System", A study paper prepared for Human Resources and Skills Development Canada (HRSDC) and Statistics Canada, Canada.
- Gagné, P., and G. St-Louis (2017), "The Statistical Registers Transformation and Integration Project at a Glance", a presentation for Advisory Committee, Statistical Registers and Geography Division, Canada: Statistics Canada.
- MacNabb, L. (2017), "Social Statistics Transformation Phase 1 Strategic Plan: Social Health and Labour Statistics Field", an internal report, Social Health and Labour Statistics Field, Canada: Statistics Canada.
- Statistics Canada, Definitions. Data sources and methods / Statistical units / Building
<http://www23.statcan.gc.ca/imdb/p3Var.pl?Function=Unit&Id=450333>
- Statistics Canada, Definitions. Data sources and methods / Statistical units / Building
<http://www23.statcan.gc.ca/imdb/p3Var.pl?Function=Unit&Id=450334>
- Statistics Canada, Illustrated glossary, Spatial Data Infrastructure (SDI)
<https://www150.statcan.gc.ca/n1/pub/92-195-x/2011001/other-autre/infrastructure-eng.htm#n1>
- Wallgren, A. and Wallgren, B. (2014), *Register-based Statistics: Statistical Methods for Administrative Data*, United Kingdom: Wiley.