

Note

« Une ressource numérique pour l'analyse historique : le recensement canadien de 1891 »

Kris Inwood et Kevin James

Cahiers québécois de démographie, vol. 34, n° 2, 2005, p. 315-328.

Pour citer cette note, utiliser l'adresse suivante :

<http://id.erudit.org/iderudit/014014ar>

Note : les règles d'écriture des références bibliographiques peuvent varier selon les différents domaines du savoir.

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter à l'URI <http://www.erudit.org/apropos/utilisation.html>

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche. Érudit offre des services d'édition numérique de documents scientifiques depuis 1998.

Pour communiquer avec les responsables d'Érudit : erudit@umontreal.ca

Une ressource numérique pour l'analyse historique : le recensement canadien de 1891

Kris INWOOD et Kevin JAMES *

À partir des recensements des 19^e et 20^e siècles, des équipes de chercheurs travaillent à la création de bases de données informatisées qui vont améliorer considérablement l'infrastructure canadienne de recherche en histoire et en sciences sociales, en démographie particulièrement. Cette note de recherche porte sur la préparation d'une bande-échantillon à grande diffusion du recensement canadien de 1891, en cours à l'Université de Guelph. Les chercheurs décrivent le projet, leur stratégie d'échantillonnage et le potentiel du nouvel outil de recherche.

English abstract, p. 328.

Les premiers recensements fédéraux réalisés au Canada, en 1871, puis en 1881 et en 1891, ont fait apparaître la vie des particuliers avec une précision et une cohérence inconnues jusque-là. Leurs informations ont beaucoup servi. Les données sur la population et l'industrie ont alimenté les débats du temps sur le succès ou l'échec de l'État canadien et de la « politique nationale » et demeurent aujourd'hui une référence indispensable à la compréhension de la société et de l'économie canadiennes.

Durant la seconde moitié du 20^e siècle, les calculatrices puis l'informatique ont engendré de nouvelles méthodes d'exploitation des anciens recensements. Avec les informations originales sur les ménages, l'agriculture et l'industrie, il est devenu possible de monter des bases de données lisibles à la machine autorisant un large éventail d'analyses. Pendant les années 1970, avec leur échantillon des ménages dénombrés au recensement canadien de 1871, Gordon Darroch et Michael Ornstein, de l'Université

* Université de Guelph. Texte traduit de l'anglais par Gilbert Gendron.

York, ont été parmi les premiers à concrétiser l'idée d'un instrument de recherche tiré d'anciens recensements, prenant la forme d'un fichier national exploitable à la machine (Darroch et Ornstein, 1979, 1980). D'autres projets se sont ajoutés à celui-là : la base de données constituée par Michael Anderson à partir du recensement mené en Angleterre et au pays de Galles en 1851 (Anderson, 1977, 1988), un échantillon (1/760) du recensement américain de 1900 réalisé à l'Université de Washington par Samuel Preston et Robert Higgs (1980) et un certain nombre de bandes-échantillons à large diffusion tirées d'anciens recensements des États-Unis (voir les comptes rendus de Ruggles, 1991, et de Ruggles et Menard, 1995).

L'importance des bandes-échantillons à grande diffusion n'a cessé de croître au cours des trois dernières décennies. Des données ainsi traitées, venant de nombreux pays, sont actuellement d'usage courant en histoire, en sciences humaines et, de plus en plus, en sciences de la santé. Elles éclairent à la fois l'aventure collective des grands changements sociaux et économiques et la diversité du vécu individuel selon les lieux et les époques. Elles permettent de comprendre l'histoire en rattachant les changements structurels et les grandes transformations à la multitude des caractéristiques et des comportements singuliers. Elles sont faciles d'accès grâce à des institutions comme le Minnesota Population Center (Ruggles, Sobek et Gardner, 1996), qui rend de précieux services aux chercheurs du monde entier : archivage, documentation, fourniture conviviale de données provenant des États-Unis, du Canada, de la Grande-Bretagne et de nombreux autres pays.

Durant les années 1980 et 1990, malgré la qualité des vieux recensements et le travail novateur de Darroch et Ornstein, l'infrastructure de recherche du Canada avait du retard en matière d'accès aux données censitaires. Une ère nouvelle a été inaugurée en 2001 par la publication d'une bande-échantillon du recensement de 1901, dans le cadre du Canadian Families Project dirigé par Eric Sager de l'Université de Victoria (Sager, 2000). Peu après, avec la collaboration du North Atlantic Population Project (Roberts et al., 2002), Lisa Dillon, de l'Université d'Ottawa et de l'Université de Montréal, a publié une base de données contenant tout le recensement de 1881 (Dillon, 2000; Dillon et Ruggles, 2001). Elle prépare actuellement un échantillon à grande diffusion du recensement de 1852. Des échantillons similaires, tirés des recensements de 1911, 1921, 1931, 1941 et 1951, sont présentement en cours de réalisation dans le cadre du vaste projet

Infrastructure de recherche sur le Canada au XX^e siècle (IRCS), sous la direction de Chad Gaffield de l'Université d'Ottawa, avec la collaboration de Statistique Canada et de chercheurs de diverses institutions.

Le rattrapage va donc bon train. En très peu de temps, le Canada se sera doté d'une remarquable série de données transversales allant de 1852 à 2001. Avec des collègues d'autres institutions, des chercheurs des départements d'histoire et d'économique de l'Université de Guelph participent à cet effort, par la création d'une base de données du recensement de 1891. Une fois terminée, la base permettra d'étudier un éventail de questions relatives aux années 1880 et 1890 et de mieux comprendre le développement social et économique du Canada depuis le 19^e siècle.

LES RICHESSES DU RECENSEMENT DE 1891

La société et l'économie canadiennes ont changé entre le début des années 1880 et la fin des années 1890, et la manière dont nous comprenons notre histoire dépend souvent du regard que nous posons sur cette évolution. Quel rôle ont joué les droits de douane, l'industrialisation, les liaisons ferroviaires transcontinentales, les mutations du monde rural, la baisse de la fécondité et la hausse de l'âge au mariage, l'immigration et l'émigration, les inégalités régionales croissantes, le boom du blé, les progrès de l'alphabétisation et autres transformations ? L'accès aux données du recensement de 1891 peut faire avancer notre connaissance de ces questions, ne serait-ce qu'en nous permettant de mieux situer les changements dans le temps, entre les points de référence que fournissaient déjà les recensements de 1881 et de 1901.

Ainsi, après la Confédération, la proportion de migrants européens accueillis au Canada est descendue à un niveau très bas pendant quelques décennies, et les natifs du Canada ont émigré en grand nombre aux États-Unis, où les salaires et les revenus étaient nettement plus intéressants. Une comparaison des indices de croissance des revenus et de l'industrialisation aux États-Unis et au Canada témoigne du retard de ce dernier (Green et Urquhart, 1993). Le vent a tourné à la fin des années 1890, avec la poussée de la colonisation dans l'ouest et la soudaine expansion économique connue sous le nom de « boom du blé » (Inwood et Stengos, 1991). Pour mieux comprendre ce mou-

vement, il faut en savoir davantage sur la réaction de la société et de l'économie canadiennes aux forces de la mondialisation et du changement technologique *avant* le boom du blé. Le recensement de 1891 est particulièrement intéressant à cet égard, étant celui qui précède le plus immédiatement ce dernier.

Autre exemple : le recensement de 1901 est riche en informations sur les immigrants (Green et MacKinnon, 2001; Green, MacKinnon et Minns, 2002), mais l'image qui en ressort gagnera en profondeur si un échantillon du recensement de 1891 nous permet d'établir leur répartition par métier dix ans auparavant et d'en déduire leurs possibilités d'ascension professionnelle, sachant que celles-ci se répercutent sur leur décision de demeurer au Canada ou d'émigrer de nouveau, vers les États-Unis. Il ne serait pas moins intéressant de savoir quelles populations vivaient dans l'ouest du Canada avant le boom du blé, et quelles étaient celles qui, dans l'est du pays, n'ont pas suivi le mouvement d'émigration. Signalons que le recensement de 1891 a été le premier à l'occasion duquel chacun a dû préciser le lieu de naissance de son père et de sa mère. Ces informations nous permettent d'embrasser les immigrants et leurs enfants nés au Canada. Elles sont précieuses, étant donné le rôle de la famille dans l'organisation sociale et économique du pays, ou l'incidence de la mobilité intergénérationnelle sur les inégalités sociales.

On s'est interrogé aussi sur l'évolution de l'industrie manufacturière après l'entrée en vigueur de la Politique nationale (1879). De qui était constituée la main-d'œuvre des nouvelles usines ? Si déjà, en 1891, les ouvriers étaient britanniques de façon aussi disproportionnée que les ouvriers qualifiés plus âgés en 1901 (Green et MacKinnon, 2001), nous serions amenés à réviser la vieille hypothèse voulant que les droits de douane aient protégé le secteur manufacturier et entraîné l'embauche d'immigrants britanniques tandis que des Canadiens émigraient aux États-Unis. Il sera possible d'éclairer cette question grâce au recensement de 1891, le premier au Canada à inclure un éventail d'informations soignées et systématiques sur l'emploi. Ces microdonnées permettront en outre de croiser les catégories professionnelles avec des variables comme l'âge, l'origine ethnique et la religion, au niveau individuel, pour vérifier si les caractéristiques de l'emploi urbain relevées par Baskerville et Sager concernent les petites villes et les régions rurales (Baskerville et Sager, 1989, 1998; Sager et Baskerville, 1990). Autre fait à souligner, les documents manuscrits de 1891 apportent, sur la codification des professions par le personnel du recensement,

une information dont l'analyse nous permettra de reconstituer la classification utilisée et, par ce moyen, d'établir si les autorités canadiennes ont « remanié » les données du recensement afin de faire disparaître le travail des femmes mariées des tableaux rendus publics, ainsi qu'on l'a découvert dans le cas des États-Unis (Carter et Sutch, 1996).

Le recensement de 1891 présente encore une particularité, qui complique l'échantillonnage mais ne manque pas d'intérêt au plan méthodologique. Cas exceptionnel après la Confédération, en 1891, les agents recenseurs n'ont pas numéroté les logements (*dwellings*); ils ont inscrit les membres des familles et attribué à celles-ci un numéro, mais sans numéroté les logements en parallèle, contrairement à la pratique observée dans les autres recensements. Il nous faut donc faire le lien par déduction, non sans risque d'erreur. Le fait que les relevés ne présentent pas systématiquement la population recensée par ménages ou par logements complique l'échantillonnage, quelle que soit la stratégie choisie (voir la section suivante), mais nous oblige à vérifier si l'information sur les immeubles est systématique. Par bonheur, ceux-ci sont décrits de manière assez détaillée. Le recensement de 1891 est le premier, après la Confédération, à relever l'information sur les matériaux, le nombre d'étages et le nombre de pièces. Ces données permettront de faire une analyse attentive des immeubles.

Le recensement de 1891 a été le premier également à prendre note du « lien avec le chef de famille »; la diffusion de cette information fera avancer la recherche sur la composition des familles et la cohabitation à cette époque (Baskerville, 2001; Darroch, 2000, 2001; Dillon, 1997). D'autre part, les données de 1891 pourraient permettre, indirectement, d'inférer plus efficacement les liens familiaux au sein de la population recensée en 1871 et en 1881 (Dillon, 1996).

Bien entendu, tout échantillon national permet aux chercheurs de mettre en contexte des collectivités ou des individus. Nous pouvons apprendre davantage d'une étude ou d'une source existante en la confrontant à un échantillon représentatif d'une société dans son ensemble. Pour ne citer qu'un exemple, Di Matteo (2001) a lié les successions homologuées de personnes décédées en 1892 à la caractérisation de ces dernières selon le recensement de 1891. Notre base de données permettra à Di Matteo d'affiner son évaluation de la représentativité des défunts par rapport à l'ensemble de la population et, par ce moyen, d'améliorer sa mesure des inégalités sociales.

Avec l'achèvement de notre travail et des autres projets similaires en cours, la série d'échantillons des recensements du Canada sera presque complète : il ne manquera plus que les années 1861 et 1961 à cette séquence d'instantanés transversaux pris à intervalles de dix ans sur une période de 150 ans. Nous pourrions bientôt mener au Canada des études semblables à celle de Ruggles et de ses collaborateurs, qui, avec une ingéniosité remarquable, ont utilisé les bases de données américaines pour dresser un tableau des changements socioéconomiques et démographiques survenus aux États-Unis en quelques générations (Fitch et Ruggles, 2000; Ruggles, 1994).

Un peu de la même façon, Bradbury (2000) relève que le pourcentage de familles monoparentales au Canada n'était pas moindre en 1901 que plus tard au cours du siècle : l'image des familles nucléaires stables de jadis serait-elle sans fondement ? En fait, en 1901, la reconfiguration des familles était causée par le veuvage précoce plus que par le divorce et la séparation. C'est ce que tendent à montrer les travaux de Céline Le Bourdais, entre autres. Les données des anciens recensements permettront de remonter le temps et de suivre le fil de cette évolution, entre changement et stabilité. Dans bien des domaines, la série longitudinale de bases de données censitaires éclairera l'évolution — ou la non-évolution — de notre pays à long terme.

Le recensement lui-même sera un objet de recherche plus accessible. Dunae (1998) s'est déjà engagé dans cette voie en étudiant les résultats du recensement de 1891 en Colombie-Britannique. L'accessibilité des anciens recensements pourra stimuler l'intérêt des chercheurs pour tout ce qui entoure la réalisation des recensements au Canada. Le bon usage des recensements repose en effet jusqu'à un certain point sur la connaissance de leur organisation et de leur déroulement sur le terrain, ainsi que du contexte culturel et politique qui les a façonnés. C'est pourquoi l'édification de notre base de données comporte la préparation de métadonnées ayant pour fonction de baliser tant la conception et l'organisation des données, que les analyses et interprétations des futurs usagers.

ÉCHANTILLONNAGE ET PROGRESSION DES TRAVAUX

Pour situer le processus d'échantillonnage dans son contexte, il faut tenir compte de la manière dont le Bureau du recensement a organisé, en 1891, le dénombrement puis le traitement et la conservation des données. Des agents recenseurs, au nombre de

4366, ont été chargés de dénombrer chacun un millier de personnes en moyenne, soit quelque 210 familles. Des commissaires de district devaient rassembler leurs relevés et les faire parvenir au Bureau du recensement, à Ottawa. Certaines inscriptions ont subi des révisions. On ne sait pas grand-chose du processus de vérification des données, mais les commissaires de district ou les fonctionnaires d'Ottawa ont vraisemblablement consulté les recenseurs dans les cas où il existait un doute sur l'information recueillie. Pour la première fois dans l'histoire des recensements canadiens, des tabulatrices ont servi à la préparation des rapports.

Après la publication des quatre volumes et des nombreux bulletins consacrés aux résultats du recensement, les relevés manuscrits des recenseurs ont été conservés par le Bureau du recensement jusqu'à leur mise sur microfilm, à la fin des années 1930. Des copies des 128 bobines de microfilm ainsi réalisées sont accessibles depuis 1983, date à laquelle elles ont été rendues publiques. Ces films contiennent notre base d'échantillonnage. Il s'agit de copies microfilmées des relevés manuscrits originaux, dont chaque page contient vingt-cinq lignes et fournit, par conséquent, des renseignements sur un maximum de vingt-cinq personnes.

La plupart des bases de données tirées des recensements nord-américains sont construites par sélection aléatoire de chefs de ménage (*dwelling heads*) ou de pages de relevés. Dans le cas des anciens recensements, la sélection par pages est généralement le procédé le moins coûteux, parce que le repérage des points d'échantillonnage est vite fait et que la numérotation des pages n'est pas une source d'imprécision autant que peut l'être la numérotation des logements ou des ménages. Pour 1891, la sélection de logements est exclue : les relevés contenaient une colonne pour la numérotation des ménages, mais non pour la numérotation des logements (et certaines familles occupaient ou possédaient plus d'un logement). Cette confusion rendrait très onéreuse la mise en œuvre d'un système de numérotation incluant tous les logements d'une province (y compris les logements non échantillonnés). Il faudrait réduire la taille de la base de données, qui perdrait ainsi de son utilité pour beaucoup d'analyses. La solution de rechange, la sélection par pages, produit des groupements de logements qui risquent également d'influencer les erreurs types. En fin de compte, il faut se demander quelle méthode a les effets les plus importants sur ces dernières.

Après avoir testé diverses stratégies d'échantillonnage sur une base de données couvrant tout le recensement américain de 1880, Ruggles et Goeken (2002) concluent à l'effet limité des groupements résultant de la sélection par pages. Il y aurait à peu près autant d'hétérogénéité dans une seule page que dans une séquence de vingt pages, comme on pouvait s'y attendre pour un recensement de porte à porte. Par conséquent, comme une simple modification de logiciel permettait de corriger les effets de groupement et que, surtout, nous ne souhaitions pas réduire la taille de l'échantillon pour des raisons de coût, la stratégie retenue a été celle de la sélection par pages. Ce choix permet d'ailleurs d'étudier les « effets de voisinage » et d'effectuer des comparaisons par paires de voisins pour neutraliser l'influence du milieu, par exemple dans l'analyse de la mortalité infantile.

Dans chaque bobine, à partir d'un point fixé au hasard, nous avons sélectionné les pages par pas de vingt, de manière à obtenir un échantillon de 5 pour cent (10 pour cent pour l'ouest du pays et les grandes villes). Notre objectif étant de capter des ménages entiers, nous avons inscrit les individus figurant sur chaque page sélectionnée en commençant par le premier nouveau logement de la page (autrement dit en éliminant les premières personnes mentionnées sur une page si le début de la page ne coïncidait pas avec le début de l'information sur un logement) et en poursuivant jusqu'à la dernière personne du dernier logement de la page (en complétant le dernier bloc d'information de la page sélectionnée à l'aide de la page suivante).

À l'étape de la saisie des données, il fallait donc repérer, pour chaque page sélectionnée, le début de l'information sur le premier logement et la fin de l'information sur le dernier logement. Le repérage de chacun des logements compris entre le premier et le dernier logement a été confié, à un stade ultérieur du traitement des données, aux programmeurs chargés de numérotter les logements. Cette façon de faire a épargné des décisions complexes au personnel chargé de l'indexation et de la saisie des données et nous a donné l'occasion de tester, plus tard dans le processus, la sensibilité de diverses méthodes de reconnaissance des logements.

Le repérage correct du premier nouveau logement d'une page sélectionnée est essentiel. Comme les familles sont numérotées, on peut dire que ce premier logement correspond à la première nouvelle famille occupant son propre logement. Le même critère, appliqué à la page suivante, permet de repérer la fin de l'information sur le dernier logement de la page sélectionnée.

Par conséquent, la page sélectionnée commence avec le début de l'information sur une famille et se termine avec la fin de l'information sur une famille. Cette stratégie permet de capter des familles entières. Notons que, dans le cadre du recensement de 1891, une famille était définie comme « des personnes qui vivent ensemble sous le même toit et dont l'approvisionnement en vivres est commun » (Canada, 1891). La plupart des chercheurs jugeront cette définition de la famille formulée en 1891 équivalente à celle du ménage dans le vocabulaire et la conceptualisation d'aujourd'hui.

Il est clair que cette stratégie d'échantillonnage exige une indexation précise et exhaustive de chacune des bobines préalablement à la sélection de l'échantillon de cinq pour cent des pages. Une lecture préliminaire de chaque bobine permet de trouver la numérotation des pages et de repérer les immeubles où vivaient plus de trente habitants. Nous avons effectué l'indexation de chaque bobine à deux reprises et résolu les divergences, puis opéré une nouvelle vérification au moyen d'un processus d'indexation supplémentaire.

Comme la plupart des échantillons de cette nature, la base de données du recensement de 1891 appliquera un traitement différent aux données sur les personnes vivant en institution ou dans des immeubles collectifs (*large dwellings*), afin de limiter les erreurs types (Ruggles, 1995). Les immeubles collectifs sont ceux où logeaient trente personnes ou davantage (hôpitaux, pensionnats, asiles, hôtels, etc.). Notre stratégie consiste à inscrire séparément tous leurs habitants, dans un premier temps, et à tirer plus tard un échantillon de cinq pour cent de cette population, en respectant les principes appliqués par le projet IRCS pour la préparation des bases de données tirées du recensement de 1911 et des recensements ultérieurs.

À terme, la base de données contiendra des informations sur environ 350 000 personnes, provenant d'échantillons de cinq pour cent de la population de l'est du Canada, de dix pour cent de celle de l'ouest du Canada, de dix pour cent de la celle des grandes villes et de tous les résidants des immeubles où vivaient plus de trente personnes. La Fondation canadienne pour l'innovation (FCI) a financé le travail sur la partie ontarienne des données, par la nomination du professeur Doug McCalla comme titulaire de la chaire de recherche du Canada sur l'histoire rurale, située à l'Université de Guelph. La saisie des données, amorcée à l'automne 2003, est terminée pour ce qui est de l'Ontario. Une seconde tranche du financement de la FCI, disponible depuis

mai 2005, nous permet d'étendre la base au reste du Canada. À terme, nous aurons créé un échantillon national intégré.

La préparation de la base de données nécessite de nombreuses opérations : indexation des 128 microfilms contenant les rapports de dénombrement, saisie des données, recherche des doubles entrées, vérification et correction, codification, construction de variables synthétiques, mise à l'essai, collecte de métadonnées contextuelles et interprétatives, traitement documentaire. Nous avons bon espoir d'avoir achevé en janvier 2007 la majeure partie sinon la totalité de la saisie des données et leur nettoyage (suréchantillons mis à part), ainsi que le traitement documentaire dans ses grandes lignes. À ce moment-là, nous pourrions mettre à la disposition des utilisateurs les fichiers de données tirés des bobines de microfilm, mais la codification ne sera vraisemblablement pas terminée. Selon nos prévisions, vers le début de 2008, il sera possible d'accéder à un grand fichier unique doté d'une codification à tout le moins partielle et contenant les numéros des logements ainsi que des suggestions de corrections aux données d'origine. Il faudra attendre l'année suivante pour effectuer un test bêta sur la version finale, qui devrait être accompagnée d'explications sur les variables, les métadonnées, la conformité aux normes de la DDI et le traitement documentaire, d'une bibliographie et de brèves présentations des données (analogues aux tableaux et commentaires placés au début de chaque volume du recensement).

Divers systèmes de classification des variables clés intégrés à la base de données permettront les comparaisons internationales et les comparaisons dans le temps (intercensitaires par exemple). On sait que des problèmes d'interprétation peuvent résulter de l'application d'un système unique de classification et de codification à des données portant sur des sociétés et des époques différentes. Cette difficulté a été résolue en grande partie par l'application d'une codification uniformisée aux données des recensements américains de 1850 à 2000 pour la réalisation des IPUMS (Integrated Public Use Microdata Series), et aux données des recensements de 1880 aux États-Unis et de 1881 au Royaume-Uni et au Canada, dans le cadre du North Atlantic Population Project. Autant que possible, nous préparerons les données de 1891 pour les rendre compatibles avec ces bases de données, et avec l'organisation et la codification d'autres bases de données canadiennes, en particulier avec les bandes-échantillons des recensements de 1911, 1921, 1931, 1941 et 1951 réalisées dans le cadre de l'IRCS.

La publication officielle des données se fera à la fin du projet, en décembre 2009. Les chercheurs qui souhaiteront avoir accès à une partie d'entre elles d'ici là seront bien accueillis, mais doivent savoir que l'état d'achèvement de ce qui leur sera rendu disponible dépendra du calendrier de réalisation de la base de données. Au terme du projet, l'échantillon principal devrait être accessible sans restrictions par le réseau de répertoires de données qui alimente le milieu canadien de la recherche historique en sciences sociales. L'Initiative de démocratisation des données, hébergée par Statistique Canada, et l'infrastructure du North Atlantic Population Project, hébergée à l'Université du Minnesota, vont vraisemblablement apporter leur concours. D'autres institutions qui diffusent des données seront invitées à acquérir des copies de la bande-échantillon et à les rendre disponibles. La publication presque simultanée de plusieurs nouvelles bases de données tirées d'anciens recensements suscitera peut-être un mouvement en faveur de la création d'archives nationales pour ces données, qu'il s'agisse d'une nouvelle institution ou d'une addition au mandat d'une institution existante (Bibliothèque et Archives Canada constitue la possibilité la plus évidente).

La réalisation de ce travail sur le recensement de 1891 aura mobilisé des collègues de plusieurs universités canadiennes, du Minnesota Population Centre et de l'U.K. Data Archive. À l'Université de Guelph, nous avons bénéficié de la collaboration des collègues des arts et des sciences sociales, de la bibliothèque, des services informatiques et de la Chaire de recherche du Canada sur l'histoire rurale.

Le projet lui-même est hébergé dans les locaux du nouveau Historical Data Centre (HDC), qui devrait lui survivre et demeurer un espace à la fois virtuel et physique où des collègues se rencontrent et unissent leurs efforts pour créer des bases de données et les utiliser pour des recherches novatrices. L'un des objectifs du projet de base de données du recensement de 1891 est en effet de susciter une collaboration plus active parmi les historiens et les spécialistes des sciences sociales de l'Université de Guelph et d'ailleurs, dans le cadre du HDC.

CONCLUSION

On consultera la page web du projet pour obtenir de plus amples renseignements sur la base de données du recensement de 1891 ¹. Il vaut la peine de rappeler que cette précieuse ressource

numérique destinée à l'analyse historique n'est pas le fruit d'une initiative isolée et que d'autres projets similaires sont en cours. L'effort actuel de mise en valeur des anciens recensements pourrait représenter le progrès le plus important de l'infrastructure de recherche en sciences humaines et sociales de l'histoire du Canada. Les données de 1891 revêtiront une utilité encore plus grande lorsqu'elles seront mises en parallèle avec les données de 1852 et de 1881 réunies par Lisa Dillon et avec les bases de données du 20^e siècle qui résulteront du projet IRCS. Nous bénéficierons collectivement de ces instruments qui vont nous aider à mieux comprendre le passé, et par conséquent le temps où nous vivons.

NOTE

- ¹ <http://www.census1891.ca/>.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ANDERSON, M. 1977. « The national sample from the 1851 census of Great Britain. An interim report on methods and procedures », *Historical Methods Newsletter*, 10, 3 : 117-121.
- ANDERSON, M. 1988 « Households, families and individuals: Some preliminary results from the national sample from the 1851 census of Great Britain », *Continuity and Change*, 3, 3 : 421-438.
- BASKERVILLE, P. 2001. « Familiar strangers. Urban families with boarders, Canada, 1901 », *Social Science History*, 25, 3 : 321-346.
- BASKERVILLE, P., et E. SAGER. 1989. « The first national unemployment survey: Unemployment and the Canadian census of 1891 », *Labour/Le travail*, 23 : 171-178.
- BASKERVILLE, P., et E. SAGER. 1998. *Unwilling Idlers: The Urban Unemployed and their Families in Late Victorian Canada*. Toronto, University of Toronto Press.
- BRADBURY, B. 2000 « Single parenthood in the past: Canadian census categories, 1891-1951 », *Historical Methods*, 33, 4 : 211-217.
- CANADA. 1891, *Manual Containing the Census Act and Instructions to Officers*. Ottawa, Department of Agriculture, Census Branch.
- CARTER, S. B., et R. SUTCH. 1996. « Fixing the facts: Editing of the 1880 U.S. census of occupations, with implications for long-term labor-force trends and the sociology of official statistics », *Historical Methods*, 29, 1 : 5-24

- DARROCH, G. 2000. « Constructing census families and classifying households: "Relationship to head of family or household" in the 1901 census of Canada », *Historical Methods*, 33, 4 : 206-210.
- DARROCH, G. 2001. « Home and away: Patterns of residence, schooling, and work among children and never married young adults, Canada, 1871 and 1901 », *Journal of Family History*, 26, 2 : 220-250.
- DARROCH, G., et M. ORNSTEIN. 1979. « Errors in historical data files », *Historical Methods*, 12, 2 : 157-167.
- DARROCH, G., et M. ORNSTEIN. 1980. « Ethnicity and occupational structure in Canada in 1871: The vertical mosaic in historical perspective », *Canadian Historical Review*, 61, 3 : 305-333.
- DILLON, L. Y. 1996. « Integrating nineteenth-century Canadian and American census data sets », *Computers and the Humanities*, 30, 5 : 381-392.
- DILLON, L. Y. 1997. *Between Generations and Across Borders: Living Arrangements of the Elderly and their Children in Victorian Canada and the United States*. University of Minnesota, thèse de doctorat inédite.
- DILLON, L. Y. 2000. « International partners, local volunteers and lots of data: The 1881 Canadian census project », *History and Computing*, 12, 2 : 163-176.
- DILLON, L. Y., et S. RUGGLES. 2001. « Creating historical snapshots of North America in 1880/1: Collaboration between historians and the Church of Jesus Christ of Latter-day Saints on the 1880/1 census databases of the United States and Canada », *Genealogical Journal*, 29, 3 : 107-113.
- DI MATTEO, L. 2001. « Patterns and determinants of wealth inequality in late-nineteenth-century Ontario », *Social Science History*, 25, 3 : 347-380.
- DUNAE, P. 1998. « Making the 1891 census in British Columbia », *Histoire sociale/Social History*, 31, 62 : 223-239.
- FITCH, C., et S. RUGGLES. 2000. « Historical trends in marriage formation », dans L. WAITE, C. BACHRACH, M. HINDIN, E. THOMSON et A. THORNTON, éd. *Ties that Bind: Perspectives on Marriage and Cohabitation*. Hawthorne, Aldine de Gruyter : 59-88.
- GREEN, A., et M. MACKINNON. 2001. « The slow assimilation of British immigrants in Canada: Evidence from Montreal and Toronto, 1901 », *Explorations in Economic History*, 38, 3 : 315-338.
- GREEN, A., M. MACKINNON et C. MINNS. 2002. « Dominion or republic? Migrants to North America from the United Kingdom, 1870-1910 », *Economic History Review*, 55, 4 : 666-696.
- GREEN, A. G., et M. C. URQUHART. 1993. « Estimates of output growth in Canada: Measurement and interpretation », dans D. McCALLA

- et M. HUBERMAN, éd. *Perspectives on Canadian Economic History*. Toronto, Copp Clark, 2^e édition : 158-176.
- INWOOD, K., et T. STENGOS, 1991. « Discontinuities in Canadian economic growth, 1870-1985 », *Explorations in Economic History*, 28, 1991, 3 : 274-286.
- ROBERTS, E., S. RUGGLES, L. DILLON, Ó. GARDARSDOTTIR, J. OLDERVOLL, G. THORVALDSEN et M. WOOLLARD. 2002. « The North Atlantic Population Project: An overview », *Historical Methods*, 36, 2 : 80-88.
- RUGGLES, S. 1991. « Comparability of the public use files of the U.S. census of population, 1880-1980 », *Social Science History*, 15, 1 : 123-158.
- RUGGLES, S. 1994. « The transformation of American family structure », *American Historical Review*, 99, 1 : 103-128.
- RUGGLES, S. 1995. « Sample designs and sampling errors », *Historical Methods*, 28, 1 : 40-42.
- RUGGLES, S., et R. GOEKEN. 2002. *Simulating Standard Errors under Different Sampling Strategies*. Communication présentée dans le cadre d'un atelier du North Atlantic Population Project, Grundarfjörður, Islande, 4 octobre.
- RUGGLES, S., et R. MENARD. 1995. « The Minnesota historical census », *Historical Methods*, 28, 1: 6-10.
- RUGGLES, S., M. SOBEK et T. GARDNER. 1996. « Disseminating historical census data on the World Wide Web », *IASSIST Quarterly*, 20, 1 : 4-18.
- SAGER, E. W. 2000. « The Canadian Families Project and the 1901 census », *Historical Methods*, 33, 4 : 179-184.
- SAGER, E. W., et P. BASKERVILLE. 1990. « Locating the unemployed in the urban British Columbia: Evidence from the 1891 census », *Journal of Canadian Studies*, 25, 3 : 38-54.

ABSTRACT

Kris INWOOD and Kevin JAMES

A DIGITAL RESOURCE FOR HISTORICAL ANALYSIS: THE 1891 CANADIAN CENSUS

Scholars are currently developing a number of historical census databases that will significantly enhance the research infrastructure for history, demography and other social sciences in Canada. One database project at the University of Guelph aims to prepare a public-use microdata sample (PUMS) from the 1891 census of Canada. This note introduces the 1891 project, describes the sampling strategy and provides some indication of the research potential of the new resource.