



**Conseil économique
et social**

Distr.
GÉNÉRALE

ECE/CES/GE.41/2007/9
22 mars 2007

FRANÇAIS
Original: ANGLAIS

COMMISSION ÉCONOMIQUE POUR L'EUROPE

CONFÉRENCE DES STATISTICIENS EUROPÉENS

Groupe d'experts des recensements de la population
et des habitations

Dixième session
Astana, 4-6 juin 2007
Point 3 b) de l'ordre du jour provisoire

**TECHNOLOGIE UTILISÉE DANS LE CADRE DES RECENSEMENTS: FAITS
NOUVEAUX ET INCIDENCES SUR LES MÉTHODES DE RECENSEMENT**

Utilisation de l'Internet pour le recensement

Communication du Canada *

Cette session est organisée conjointement avec Eurostat.

Résumé

Le Bureau de la Conférence des statisticiens européens a approuvé, lors de la réunion qu'il a tenue à Washington D.C. (États-Unis) les 19 et 20 octobre 2006, le renouvellement du mandat du Groupe directeur pour les recensements de la population et des habitations ainsi que le plan des futures activités de la Conférence concernant les recensements de la population et des habitations. Le Bureau est également convenu que le Groupe directeur coordonnerait les travaux concernant les divers types de réunions. Le présent document a été établi à la demande du Groupe directeur afin d'être présenté et examiné à la Réunion commune CEE/Eurostat sur les recensements de la population et des habitations qui aura lieu à Astana (Kazakhstan) du 4 au 6 juin 2007. Il constitue une base concrète de discussion pour la séance qui sera consacrée à la «technologie utilisée dans le cadre des recensements: faits nouveaux et incidences sur les méthodes de recensement».

* Document établi par Statistique Canada.

I. INTRODUCTION

1. Statistique Canada, comme plusieurs autres organismes de statistique dans le monde, a offert une application en ligne sécurisée sur l'Internet afin que toute personne vivant au Canada puisse remplir en ligne le questionnaire du recensement de 2006. Le présent document récapitule de manière succincte un certain nombre d'enseignements que Statistique Canada a dégagés à la suite de la décision qui avait été prise de proposer une option de réponse en ligne pour 2006 et de soumettre cette option à une «répétition générale» de relativement grande envergure en mai 2004. L'introduction de ce nouveau mode de réponse est l'un des principaux changements apportés dans les méthodes de recensement appliquées au Canada depuis plus de trente ans. Plus précisément, le présent document analyse les conditions de base qui doivent être réunies avant qu'il soit possible d'envisager de proposer une option de recensement en ligne, les considérations d'ordre stratégique et économique pertinentes, les principaux risques, les difficultés et les avantages de cette formule, ainsi que les perspectives qu'elle ouvre à long terme.

II. RAPPEL DES FAITS

2. Dans le cadre du recensement de 2006, les répondants au sein de l'ensemble des ménages ordinaires et des exploitations agricoles du Canada ont eu la faculté de remplir soit la version complète (53 questions), soit la version abrégée (8 questions) du questionnaire par le biais de l'Internet. Il était prévu que 16 % environ des répondants se serviraient de cette application en ligne, qui était disponible 24 heures sur 24 et 7 jours sur 7. Une vaste stratégie d'information du public devait encourager les répondants à remplir le formulaire de recensement en ligne. Des questionnaires ont été envoyés par la poste à quelque 73 % des logements au Canada, et le reste a été remis de la main à la main par un agent recenseur. Un code d'accès Internet unique était imprimé sur la couverture du questionnaire papier. Les répondants ont rempli leur questionnaire et l'ont transmis directement au centre de traitement des données de recensement de Statistique Canada, où leur réponse a été enregistrée et où les données qu'ils avaient fournies ont été incorporées dans le flux ordinaire des données issues du recensement. Les réponses reçues sur papier et par voie électronique ont été inscrites en regard d'une liste générale de contrôle des logements afin de déterminer ceux pour lesquels aucune réponse n'avait été donnée. Cette procédure était essentielle pour pouvoir indiquer aux agents de terrain s'ils devaient ou non procéder à un suivi des cas de non-réponse.

3. L'application Internet était conçue de façon à ne laisser aucune trace (empreinte) sur l'ordinateur du répondant une fois que celui-ci avait renvoyé le questionnaire. Les répondants qui remplissaient la version complète du questionnaire pouvaient sauvegarder leur formulaire rempli en partie et créer un mot de passe. En utilisant simultanément leur mot de passe et leur code d'accès Internet, ils pouvaient remplir leur formulaire en plusieurs sessions. Si un répondant décidait de sauvegarder son formulaire et ne le renvoyait pas dans les délais prescrits, le formulaire partiellement rempli était automatiquement renvoyé sous son nom et passait ensuite par les différentes étapes de traitement. En outre, les répondants en ligne qui sauvegardaient leur formulaire pouvaient y avoir accès à partir de différents endroits (formule itinérante). Par exemple, ils pouvaient commencer à remplir le formulaire chez eux, le sauvegarder alors qu'il était rempli en partie et finir de répondre aux questions sur leur lieu de travail. L'application Internet était disponible dans les deux langues officielles du Canada (anglais et français), et les répondants pouvaient passer d'une langue à l'autre à mesure qu'ils remplissaient

le formulaire. Une fois le questionnaire en ligne envoyé, son traitement ultérieur était indépendant du mode de réponse, que ce soit un formulaire papier ou une interview téléphonique assistée par ordinateur; il s'agissait donc d'une démarche véritablement intégrée.

III. CONDITIONS PRÉALABLES

4. D'après l'expérience acquise au Canada, un certain nombre de facteurs doivent être considérés comme indispensables avant qu'une option de recensement par l'Internet soit offerte à la population d'un pays. En outre, plusieurs éléments déterminants seraient de nature, s'ils sont présents, à accroître notablement la probabilité que cette option donne de relativement bons résultats.

A. Une méthode qui favorise l'adjonction d'une option Internet

5. Pour pouvoir proposer une option Internet, il est absolument indispensable qu'une grande partie de la population maîtrise suffisamment l'outil informatique et soit capable d'un autorecensement. Bien que le mode de distribution du questionnaire (manuel ou par voie postale) ne soit pas déterminant, il est par contre essentiel qu'un code d'accès Internet unique associé à une adresse géographique soit attribué à chaque ménage (ou individu) afin qu'il soit possible de mettre en place un système de contrôle et de suivi efficace. L'emploi systématique ou généralisé d'une méthode de recensement direct (entretien face à face) poserait à n'en pas douter un certain nombre de problèmes non négligeables pour ce qui est de la viabilité d'une option en ligne. Pour le recensement de 2006, Statistique Canada a créé une liste (le «Système principal de contrôle», dit MCS) de tous les logements auxquels a été attribué un identificateur unique associé à une adresse de voirie (celle à laquelle un questionnaire a été expédié par la poste) ou à un code géographique (correspondant à l'adresse à laquelle le questionnaire a été manuellement déposé). Cet identificateur unique et le code d'accès Internet unique étaient préimprimés sur chaque questionnaire. Les questionnaires renvoyés par voie postale, reçus par le biais de l'Internet ou encore remplis à la faveur d'une interview téléphonique assistée par ordinateur (ITAO) ont été enregistrés avec cet identificateur unique, et le MCS a été actualisé. Des listes de suivi en cas de non-réponse ont été établies à partir du MCS et communiquées périodiquement aux agents de terrain, afin qu'ils prennent contact uniquement avec les répondants correspondant aux logements pour lesquels aucune réponse n'avait été reçue.

B. Connectivité et attentes du public

6. À l'évidence, tout pays qui étudie la viabilité d'une formule de réponse en ligne à un questionnaire de recensement doit prendre en compte le degré de connectivité au sein de la population, ses possibilités d'accès à l'Internet «grande vitesse» ainsi que le degré et la facilité d'utilisation de l'Internet par la population pour ses diverses activités (opérations bancaires en ligne, déclaration d'impôt en ligne, etc.). Ce dernier élément peut être révélateur du degré de confiance de la population dans les applications en ligne et donner une idée de la mesure dans laquelle les répondants sont prêts à fournir en ligne des données de recensement confidentielles les concernant.

7. Au Canada, la majorité des ménages est reliée à l'Internet. D'après les estimations tirées de l'enquête de 2003 sur l'utilisation de l'Internet par les ménages, 7,9 millions (soit 64 %) des 12,3 millions de ménages au Canada comptaient au moins un membre utilisant régulièrement

l'Internet en 2003, à domicile, au travail, à l'école, dans une bibliothèque publique ou ailleurs. Selon les estimations, 6,7 millions de ménages (soit 54 %) comptaient au moins un membre utilisant régulièrement l'Internet à domicile, dont 4,4 millions (soit 65 % des utilisateurs à domicile) avaient accès à l'Internet grande vitesse. En outre, environ 57 % des ménages utilisant l'Internet à domicile comptaient un membre qui avait accès à des services bancaires en ligne, soit un taux bien supérieur aux 44 % enregistrés en 2001. Cette progression montre que les Canadiens ont de plus en plus confiance dans la sécurité offerte par l'Internet.

8. La proposition d'offrir une option en ligne a été largement déterminée par le fait que la très forte pénétration de l'Internet dans les ménages canadiens ainsi que son utilisation complexe pour différents types de transaction allaient susciter dans la population une solide aspiration à se voir offrir une option en ligne sécurisée et efficace pour le recensement de 2006.

C. Une infrastructure de sécurité bien développée

9. S'il est vrai que les Canadiens en général prennent goût à l'Internet, leur confiance dans le degré de sécurité assuré par le vendeur ou l'institution joue un grand rôle dans leur décision d'effectuer ou non une transaction en ligne. Pour un organisme de statistique, il est donc absolument essentiel de mettre en place une infrastructure sécurisée afin que les répondants aient suffisamment confiance dans le système pour accepter de communiquer en ligne des données de recensement confidentielles. Pour Statistique Canada, étant donné qu'il s'agissait d'une première offre à l'échelon national, il était important de mettre en place une infrastructure qui soit en réalité aussi, voire plus, sûre que la plupart des autres systèmes en ligne auxquels les Canadiens étaient devenus habitués. La différence fondamentale, dans le cas d'une enquête statistique ou d'un recensement, tient au fait qu'il est demandé à l'enquêté ou au recensé d'apporter des réponses à des questions, et non de prendre l'initiative ou de bénéficier directement d'une transaction, comme dans le cas d'une opération bancaire en ligne ou d'une demande d'allocation-chômage.

10. Il faut étudier attentivement le degré et le type de sécurité recherché car ces considérations influent sur la conception et le coût global de l'option Internet. Il existe toute une gamme d'options, et elle va certainement s'élargir dans les années à venir. Au Canada, l'une des solutions consistait à mettre au point une application téléchargeable qui aurait permis au répondant de remplir un formulaire et aurait chiffré ensuite les données fournies avant qu'elles ne soient transmises. C'était là le modèle de petite envergure qui avait été mis à l'essai en 2001; en raison de la durée du téléchargement et d'autres facteurs techniques, il a été décidé de procéder autrement en 2006. Une autre solution consistait à utiliser le SSL128 dont se servent un certain nombre de banques. Cette solution offre un degré de sécurité relativement satisfaisant mais Statistique Canada a estimé qu'il était nécessaire, pour le recensement, d'utiliser une formule encore plus sûre.

11. Dans le cadre de sa stratégie «Gouvernement en direct» (GED) qui prévoyait que tous les services devaient être informatisés pour 2005, le Gouvernement canadien a créé une «voie de communication protégée» qui permettait aux citoyens de transmettre des renseignements confidentiels par l'Internet. Les ministères fédéraux se sont vu offrir la possibilité de l'utiliser pour appliquer leur propre stratégie aux fins de la mise en place de services en ligne. À l'origine, la voie de communication protégée devait servir uniquement pour un chiffrement unidirectionnel (c'est-à-dire un chiffrement des données transmises de l'ordinateur de l'utilisateur aux serveurs

d'un ministère). Statistique Canada avait besoin d'un chiffrement bidirectionnel étant donné que des renseignements confidentiels seraient également transmis depuis ses serveurs jusqu'à l'ordinateur de l'utilisateur. Cette fonctionnalité était nécessaire afin de procéder à des vérifications en ligne et d'offrir aux répondants la possibilité de sauvegarder des données et de continuer à remplir le formulaire par la suite. Statistique Canada a réalisé l'investissement nécessaire pour adapter la voie de communication protégée en fonction des besoins du recensement. Cette adaptation, associée à l'utilisation de certificats anonymes à usage limité et à l'isolement du réseau interne de Statistique Canada, a permis une sécurisation particulièrement élevée des informations. Plus précisément, Statistique Canada voulait exclure tout risque, et dissiper toute impression qu'il puisse exister un risque, que les données de recensement communiquées par les Canadiens tombent dans d'autres mains que celles du personnel de Statistique Canada et s'assurer que les certificats numériques n'avaient aucun caractère «générique» mais servaient exclusivement pour le recensement.

12. L'échelle et la portée du recensement exigent une solide infrastructure de sécurité, et même si Statistique Canada avait les moyens de la mettre en place pour le seul recensement, grâce à la possibilité d'adapter l'initiative gouvernementale pour répondre aux besoins tout à fait particuliers de Statistique Canada, il a été jugé moins risqué et plus rentable de tirer parti des investissements consacrés à cette initiative.

D. Expérience tirée des enquêtes en ligne

13. Il est ressorti des consultations qui ont précédé le recensement de 2001 qu'une partie croissante de la population souhaitait et escomptait pouvoir utiliser une application en ligne pour le recensement de 2001. Statistique Canada a répondu à cette demande en offrant une application relativement inefficace mais sûre et fonctionnelle à un très petit sous-ensemble de la population canadienne. L'option Internet a été mise à la disposition de quelque 190 000 ménages et 8 300 exploitations agricoles sur deux sites expérimentaux bien délimités. La seule publicité donnée à cette option consistait en un encart inclus dans le questionnaire, lequel contenait aussi le code d'accès unique attribué au répondant. Pour des raisons de sécurité, l'application en ligne devait être téléchargée sur l'ordinateur du répondant, ce qui s'est avéré onéreux car l'application représentait entre 6 et 7 MB, soit un temps de téléchargement d'environ trente minutes avec un modem de 56 K. Au total, 4 300 réponses ont été reçues par l'Internet. Les enseignements tirés de cette expérience portaient principalement sur la conception de l'application ainsi que sur les recommandations d'offrir pour 2006 une véritable application en ligne qui ne devait pas être téléchargée et en fait ne laisserait aucune trace sur l'ordinateur de l'utilisateur.

14. Ne pas offrir une option Internet en 2006, considérant la demande et les attentes exprimées en 2001, aurait été contraire à l'initiative GED du Gouvernement. L'absence d'option Internet risquait d'entraîner un certain nombre de non-réponses dans des groupes qui espéraient et voulaient se voir offrir une option de réponse en ligne, ce qui faisait planer la menace d'un risque et de coûts inacceptables pour 2011. Statistique Canada a fait appel aux compétences du secteur privé pour mettre au point l'application/infrastructure nécessaire pour le recensement de 2006. Nonobstant la contribution du secteur privé, il est avantageux de pouvoir exploiter l'expérience et les connaissances disponibles en interne pour définir, expliciter et contrôler les besoins des usagers dans un contexte d'enquête.

15. Statistique Canada a offert aux répondants la possibilité de fournir des données par voie électronique lors d'un grand nombre d'enquêtes auprès des entreprises. Le but recherché consistait dans une large mesure à réduire le travail à la charge des entreprises qui fournissaient régulièrement des données, et il existe de notables différences entre le recensement et ces enquêtes, s'agissant aussi bien de leur contenu que de la population visée, mais Statistique Canada est parvenu à tirer parti des possibilités qu'il avait offertes pour planifier et mettre en place l'infrastructure de recensement ainsi que pour éliminer un certain nombre de risques.

16. Ainsi, avant de décider d'offrir une option de réponse en ligne pour les recensements, il est jugé essentiel de s'assurer de l'existence d'un certain nombre d'éléments, par exemple une méthode de collecte compatible, un système de suivi et de contrôle efficace, un taux de connectivité élevé et l'acceptation par la population ainsi que la présence d'une infrastructure de sécurité bien maîtrisée. L'expérience tirée des enquêtes et/ou recensements précédents en ligne est naturellement considérée comme un atout souhaitable pour parvenir à réduire un certain nombre de risques associés à une telle entreprise. Plus importante encore que tout autre facteur, cependant, est la volonté manifestée par l'organisme de statistique et tout organisme de financement d'envisager une option Internet comme un investissement potentiel sur le long terme tout en étant bien conscient des risques potentiels et des perspectives possibles. Il faut manifester s'assurer à l'avance que des fonds suffisants seront disponibles et ne pas compter sur des économies ultérieures aléatoires.

IV. CONSIDÉRATIONS D'ORDRE STRATÉGIQUE

17. Mis à part les conditions préalables essentielles et hautement souhaitables pour mettre en place une option de réponse en ligne pour le recensement, un certain nombre de considérations d'ordre stratégique jouent un grand rôle dans la manière de concevoir, planifier et appliquer ce genre de formule.

A. Investissements et économies potentielles

18. Les principaux facteurs de coût qui entrent en jeu dans l'application Internet sont les dépenses concernant l'élaboration et l'expérimentation des questionnaires, la mise en place d'une infrastructure sécurisée, la mise au point de l'application, les essais (fonctionnels, intégrés et avec des données réelles) ainsi que les services de soutien (appui technique et aide aux répondants), l'infrastructure de télécommunications, le matériel et les logiciels, ainsi que l'information du public. Dans le cadre du recensement de 2006, la moitié environ des dépenses était liée à l'adaptation de la voie de communication protégée et pouvait éventuellement être considérée comme un investissement ponctuel. Le nombre maximal de personnes utilisant simultanément le système est un facteur décisif pour en déterminer la capacité et par conséquent le coût; en l'occurrence, l'infrastructure devait pouvoir absorber 15 000 sessions d'utilisateurs simultanées. Le taux de réponse total par l'Internet choisi comme objectif était de 20 % des ménages avant les opérations de suivi des cas de non-réponse (soit environ 2 millions de questionnaires); si cet objectif n'était pas atteint, les conséquences s'en feraient directement sentir sur la capacité des systèmes de manutention des documents et sur les coûts de fonctionnement. En 2006, avant le début des opérations de suivi des cas de non-réponse, environ 22 % des ménages avaient répondu par le biais de l'Internet. Étant donné que ces opérations de suivi sont presque exclusivement réalisées à l'aide de questionnaires papier, quelque

2,26 millions de ménages au total (soit 18,5 %) avaient utilisé l'Internet au Canada pour répondre au questionnaire à l'issue des opérations de collecte des données.

19. Les objectifs étaient fixés pour toute la période de collecte des données de recensement (soit quelque neuf semaines en tout), mais Statistique Canada devait spécifier et mettre en place les moyens techniques permettant de traiter un nombre déterminé de sessions à tout moment. Avec très peu d'expérience et disposant seulement d'un modèle théorique sur lequel s'appuyer pour estimer l'infrastructure nécessaire, Statistique Canada a mis au point un système d'étroite surveillance, de contrôle et d'intensification et a pris l'initiative en cas de surcharge d'afficher un message demandant de différer la session afin de maintenir un équilibre entre la charge, l'exécution, le coût et le risque potentiel. Cet affichage permettait à un nombre prédéterminé de répondants déjà connectés au système de continuer à remplir le formulaire sans interruption à un rythme acceptable et empêchait d'autres utilisateurs d'accéder au système dès lors que l'infrastructure globale ne pouvait en accueillir davantage. Statistique Canada a dû faire face à une énorme demande la nuit du recensement, en particulier entre 18 et 23 heures, et a utilisé ce message à diverses reprises pendant cette période de pointe. Selon les estimations, il a ainsi été demandé à environ 150 000 utilisateurs de réessayer ultérieurement.

20. Les économies susceptibles d'être réalisées sont dues à la réduction des frais de manutention des documents, des frais de renvoi des questionnaires par la poste et des dépenses de suivi en cas de lacunes ou d'incohérence dans les réponses. Pour 2006, avec un taux de réponse par l'Internet de 16 %, ces économies devaient être à peu près du même ordre que les dépenses de mise au point et de gestion de l'application Internet, ce qui correspondait donc pratiquement à un point d'équilibre. Cet objectif a été atteint voire légèrement dépassé pendant le recensement de 2006.

21. S'appuyant sur la volonté manifestée par les Canadiens d'utiliser l'application en ligne pour le recensement de 2006, Statistique Canada est en train de formuler pour 2011 des hypothèses de planification, afin d'accroître au maximum les avantages que la réponse en ligne procure aux répondants et de faire faire une économie de coûts à l'administration.

B. Mise au point de l'application en interne ou recours à des services extérieurs?

22. En 2001, avec le concours de conseillers internes et externes, Statistique Canada a défini des critères et les a appliqués par la suite pour déterminer les composantes du recensement de 2006 qu'il valait mieux mettre au point en interne plutôt que de faire appel au secteur privé. L'une des principales considérations retenues concernait le niveau de compétences techniques et le contexte dans lequel ces compétences pouvaient être exploitées au mieux afin de réduire autant que faire se peut les risques liés à l'élaboration de l'application, à son intégration et à sa mise en œuvre. Statistique Canada est arrivé à la conclusion que ses services possédaient les compétences nécessaires pour concevoir, expérimenter avec les utilisateurs et définir le format du questionnaire en ligne, le flux des données, les aiguillages automatisés et les règles en matière de vérification en ligne. Ces éléments ont une incidence sur la manière dont les répondants apprécient l'application, sur l'allure générale et l'attrait de l'application, sur la comparabilité du questionnaire Internet avec le questionnaire papier (effets modaux éventuels) ainsi que sur la qualité des données et, comme tels, ils ont été jugés essentiels par les spécialistes de l'organisme. En revanche, le secteur privé était le mieux à même de résoudre les difficultés d'intégration des systèmes, associées à la programmation de l'application en ligne et à sa mise en relation avec les

systèmes de traitement en aval, ainsi que de s'assurer qu'elle était conçue de manière optimale pour traiter les volumes de données prévus.

23. Par conséquent, établir un profil de compétence associé à une vue d'ensemble du système général de recensement et de l'organisation des processus s'est avéré bénéfique pour exploiter au mieux les talents du personnel en interne et ceux des spécialistes du secteur privé.

24. Cette répartition des domaines de compétence s'est faite à un niveau élevé mais il existe des zones d'ombre où les capacités se chevauchent, et Statistique Canada a travaillé de concert avec des sous-traitants extérieurs afin de mettre au point l'application Internet pour le recensement dans un esprit d'équipe. Ce travail en coopération a porté ses fruits, même dans les domaines où Statistique Canada a généralement l'avantage. Les organismes de statistique sont de loin les plus qualifiés lorsqu'il s'agit d'élaborer des questionnaires papier. Des décennies d'expérience, des échantillons de très grande taille, des analyses exhaustives de données, ainsi que des essais et études d'ordre qualitatif leur ont permis d'atteindre un niveau de compétence sans égal dans le secteur privé. Statistique Canada a une expérience plus limitée en ce qui concerne la création d'outils de collecte de données sur le Web et, dans certains cas, en cherchant à suivre de trop près le formulaire papier existant dans l'espoir de réduire l'effet modal au minimum, il risque en fait de créer un effet modal encore plus important. Les répondants qui choisissent l'option Internet sont généralement des utilisateurs expérimentés de l'Internet et ont certaines attentes quant à la façon de remplir le questionnaire, attentes qui sont fonction de leur expérience d'autres formulaires électroniques. Une collaboration étroite avec les sous-traitants ainsi que le test de recensement ont permis de dégager des enseignements utiles quant à la manière de mieux concevoir le questionnaire de recensement par l'Internet afin d'obtenir les résultats souhaités. Par exemple, dans les formulaires électroniques, l'ordre des questions suit généralement la séquence oui/non de façon systématique. Dans les questionnaires papier, la réponse la plus souvent sélectionnée apparaît généralement en premier. Sur le Web, les utilisateurs retiennent apparemment la séquence oui/non dès la première fois qu'ils ont à répondre à ce type de questions, et s'attendent ensuite à ce que la séquence soit la même pour toutes les autres questions. Inverser l'ordre des réponses sur le Web afin de l'aligner sur celui du formulaire papier risque de produire l'effet contraire à celui souhaité.

25. Un autre exemple concerne les questions pour lesquelles il n'y a qu'une seule réponse logique (par exemple homme/femme) cochée par un bouton radio. Si une utilisatrice sélectionne «homme» par mégarde puis rectifie son erreur en sélectionnant «femme», la première option est automatiquement désélectionnée. Un utilisateur expérimenté de l'Internet s'attend à cet effacement automatique et le risque de se retrouver avec une double réponse «homme» et «femme» est plus grand si des boutons radio ne sont pas prévus au moment de la conception du questionnaire. Au Canada, avant d'introduire dans le formulaire Internet quoi que ce soit de différent par rapport au formulaire papier, on en a étudié avec soin les conséquences potentielles et, pour 2006, il a été possible de parvenir à un compromis entre la concordance globale avec le formulaire papier et l'incorporation de nombreuses normes utiles caractéristiques du Web.

26. Enfin, malgré les mesures de sécurité rigoureuses mises en place pour protéger les données confidentielles, il faut considérer avec attention l'inquiétude réelle ou imaginaire que peuvent manifester les répondants au sujet du caractère confidentiel et de la protection des données en apprenant la participation du secteur privé. Toute réaction négative généralisée pourrait avoir un

impact négatif sur les taux de réponse globaux et rendre difficiles les relations entre le répondant et l'organisme de statistique.

V. PRINCIPAUX RISQUES ET ENJEUX

A. Intégration

27. L'un des principaux problèmes d'intégration consistait à déterminer comment fusionner les données transmises par l'Internet et celles reçues par d'autres voies, par exemple sur papier et par téléphone. Notre décision d'appliquer certaines normes du Web a entraîné un petit nombre de différences entre les données reçues par l'Internet et celles recueillies sur papier. Il a donc fallu instaurer une coopération entre les responsables de la création de la base de données et les concepteurs du système de saisie des données sur papier et du questionnaire Internet à toutes les étapes du cycle d'élaboration.

28. Pour tenter de retirer le meilleur profit possible de l'élaboration de l'application Internet, il a été volontairement décidé d'entrée de jeu de s'efforcer d'utiliser la même logique de base et le même jeu de codets pour deux autres applications analogues, en n'y apportant que les modifications nécessaires. Plus précisément, dans ce contexte, des questionnaires ont été remplis par téléphone en mode ITAO (interview téléphonique assistée par ordinateur) et une application analogue a été utilisée pour le suivi des rejets au contrôle (FEFU) là encore en mode ITAO. Ces trois applications (Internet, ITAO et FEFU) partageaient le même jeu de codets de base, avec quelques adaptations. Cette formule a donné quelques bons résultats mais a entraîné un certain nombre de difficultés. Par exemple, l'application Internet comprenait une logique de vérification intégrée, mais l'application FEFU devait également traiter des données recueillies avec les questionnaires papier, auquel cas les aiguillages n'étaient pas nécessairement respectés et des réponses multiples inappropriées pouvaient avoir été fournies. Dans l'application FEFU, l'opérateur devait s'efforcer de faire disparaître certaines incohérences avec le concours du répondant, en étant à même d'accéder à des réponses qui ne respectaient pas la logique de l'application Internet. Ce genre de questions a entraîné un certain nombre de difficultés lorsqu'il s'est agi d'adapter l'application FEFU pour le recensement de 2006. Il est actuellement prévu que l'application FEFU et l'application Internet n'aient pas le même jeu de codets de base pour 2011.

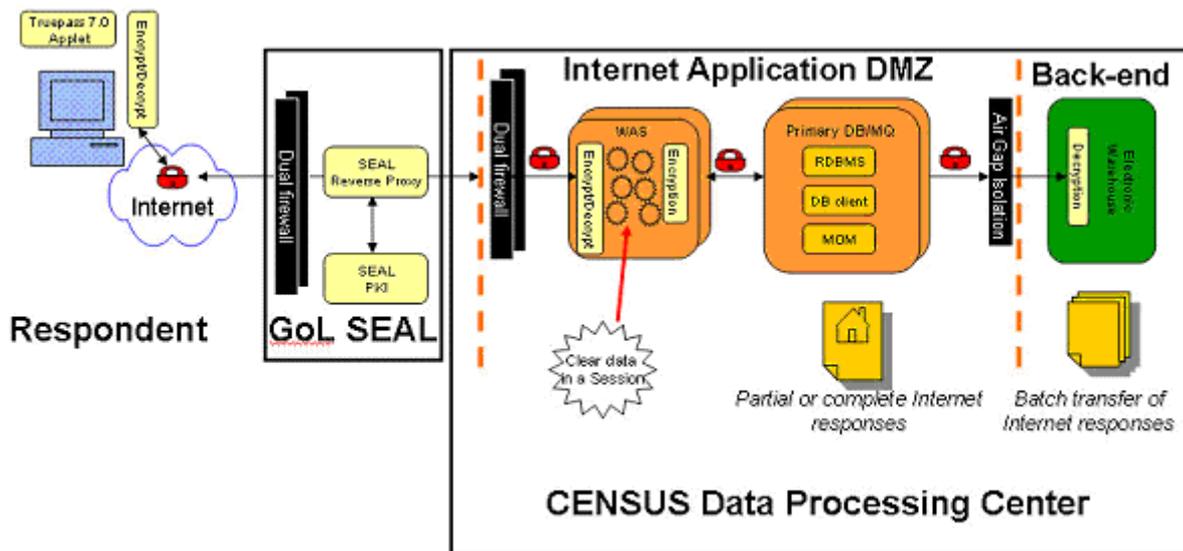
B. Infrastructure

29. L'une des principales difficultés rencontrées pour appliquer une option de réponse en ligne pour le recensement tenait à la taille de l'infrastructure nécessaire pour absorber le nombre d'utilisateurs en période de pointe. L'infrastructure comprend généralement des serveurs, des dispositifs de stockage et des moyens de communication qu'il peut être extrêmement coûteux d'acquérir, d'installer et d'essayer. La stratégie consistait à trouver un équilibre entre ce qui constituait une dépense raisonnable et le risque de ne pas pouvoir absorber le nombre d'utilisateurs simultanés en période de pointe et de les amener à abandonner l'option Internet pour remplir des questionnaires papier ou, au pire, à renoncer purement et simplement à remplir le questionnaire. La charge maximale imposée à l'infrastructure dépend de trois variables: le nombre d'utilisateurs qui se connectent à un moment donné pendant la période de pointe, le nombre moyen de pages à visualiser par chaque utilisateur et la durée moyenne de visualisation de chaque page. Par exemple, un scénario dans lequel 50 utilisateurs visualisent

chacun une page pendant cinq minutes nécessite une infrastructure moins solide que celui prévoyant 15 000 utilisateurs visionnant 20 pages à raison de dix secondes en moyenne par page.

30. Le calcul servant à déterminer l'infrastructure nécessaire fait également intervenir le nombre total de formulaires en ligne attendus. Toutefois, si les utilisateurs qui se connectent et qui envoient des formulaires représentaient la charge la plus lourde que devait supporter l'infrastructure, il fallait également prendre en compte d'autres types d'activités qui absorbent des ressources du système. Il s'agit notamment des utilisateurs qui cherchent à se connecter mais utilisent des codes d'accès non valables, des utilisateurs qui cessent de remplir leur formulaire et le sauvegardent afin de le reprendre un autre jour et de ceux qui se connectent, commencent à remplir leur formulaire mais omettent de le sauvegarder ou de l'envoyer. Les hypothèses retenues pour établir une estimation de l'infrastructure nécessaire doivent prendre en compte des estimations concernant ces activités, en plus du nombre de formulaires qui seront envoyés.

31. Un autre élément important qui entre en jeu dans le choix de l'infrastructure concerne l'architecture du système, la manière dont elle s'articule avec les autres composantes du système de traitement, et la façon dont elle sécurise les données communiquées par les répondants. Les systèmes de traitement de Statistique Canada ne sont pas raccordés à des réseaux extérieurs. Le système de réponse par l'Internet (IRC) est naturellement relié au Web. L'architecture de Statistique Canada protège les données des répondants stockées dans les serveurs de l'IRC grâce à un double pare-feu et un chiffrement des données fournies par les répondants. Elle permet aussi de transférer les données chiffrées provenant des formulaires qui ont été renvoyés vers les principaux systèmes de traitement en utilisant un dispositif d'isolement qui sépare physiquement les deux réseaux tout en permettant le passage des données autorisées à travers un commutateur, à intervalles prédéterminés.



C. Communications

32. Pour avoir une chance raisonnable d'obtenir le taux d'utilisation souhaité, on a élaboré un message faisant systématiquement état de la facilité d'utilisation de l'Internet, qui a été utilisé sur tous les supports, dans toute la publicité et dans tous les entretiens avec les médias.

33. Voici quelques exemples de la façon dont un tel message a été utilisé et du contexte dans lequel il l'a été:

- a) Tous les documents imprimés et électroniques comportaient une adresse Internet et le message «Remplissez votre questionnaire de recensement en ligne aujourd'hui»;
- b) Les éléments graphiques utilisés pour l'affiche du recensement national ont été testés auprès de groupes cibles pour vérifier que l'image et le message étaient faciles à comprendre;
- c) La publicité radiophonique payante insistait sur la facilité à répondre au questionnaire du recensement en ligne. La publicité imprimée et la publicité extérieure reprenaient l'image de l'enveloppe et de la souris;
- d) Toutes les actions de communication faisaient une large place à l'option Internet pour 2006;
- e) Des centaines d'organismes publics et privés ont accepté de faire la promotion de l'option en ligne sur leur site Web extérieur pendant les six semaines précédant le jour du recensement, en associant informations et liens hypertexte au questionnaire en ligne;
- f) Des liens avec le questionnaire en ligne ont été indiqués de façon bien visible sur le site de Statistique Canada;
- g) Le personnel de Statistique Canada a préconisé l'utilisation du questionnaire en ligne à maintes reprises au cours des interviews dans la presse ou sur les chaînes de radio ou de télévision.

34. L'incitation à utiliser l'option en ligne a posé quelques difficultés tout à fait particulières. Par exemple:

- a) L'activité sur le Web a été inopinément forte le jour du recensement, ce qui a entraîné des temps d'attente pour accéder au formulaire en ligne;
- b) Dans un premier temps, les utilisateurs de Linux ne pouvaient utiliser le questionnaire en ligne. Ce problème technique a rapidement été résolu par Statistique Canada.

VI. PRINCIPAUX AVANTAGES

A. Qualité

35. D'après un certain nombre d'études, les réponses reçues par l'Internet sont plus complètes que celles fournies sur papier. Lors du test de recensement, on a constaté que le taux de non-réponse partielle était bien plus faible dans le cas des questionnaires renvoyés par l'Internet que dans celui des questionnaires papier. Plus précisément, s'agissant de la version abrégée du formulaire de recensement, il était de 0,01 % pour les réponses communiquées par l'Internet et de 2,54 % pour les réponses fournies sur papier. Dans le cas de la version complète, il était de 1,80 % pour les réponses reçues par l'Internet et de 6,97 % pour celles fournies sur papier.

36. Ces résultats peuvent s'expliquer en partie par le fait que l'utilisateur moyen de l'Internet fournit généralement des réponses plus complètes que celui qui utilise des documents papier. Pour déterminer si l'option Internet a amélioré la qualité déjà élevée des réponses soumises par l'ensemble des utilisateurs de l'Internet, on a comparé les données fournies par chacun des répondants au test de recensement de 2004 avec celles qu'il avait communiquées sur papier pour le recensement de 2001. Lorsque l'on a pris en compte le fait que le test de recensement de 2004 avait été réalisé sur la base du volontariat et que le recensement de 2001 avait un caractère obligatoire, on a constaté que les questionnaires qu'ils avaient renvoyés par l'Internet en 2004 étaient plus complets, et présentaient un taux de non-réponse partielle plus faible que les questionnaires renvoyés en 2001.

37. Plusieurs facteurs contribuent à l'amélioration de la qualité observée dans le cas des réponses en ligne. Un certain nombre de vérifications en ligne ont été délibérément intégrées dans l'application pour alerter les répondants lorsqu'ils n'avaient pas répondu à certaines questions ou avaient enregistré des informations incompatibles entre elles. Des aiguillages automatisés ont été programmés afin d'indiquer aux répondants de sauter les questions sans objet pour eux, et les questions ont été personnalisées avec l'indication du nom du répondant, ce qui diminue le risque que les réponses d'une personne soient par mégarde enregistrées dans la zone de réponse d'une autre personne. En outre, aucune erreur de saisie des données n'intervient pendant le traitement. L'utilisation de boutons radio pour certaines questions élimine la possibilité de donner des réponses contradictoires, et celle de menus déroulants, le cas échéant, aide les utilisateurs à communiquer des réponses appropriées. Enfin, il se dégage chez les utilisateurs une impression générale, confirmée par une enquête réalisée auprès d'un échantillon de répondants par l'Internet après le test de recensement, que le formulaire Internet est facile à utiliser et demande peu de temps pour être rempli. C'est sans doute, en partie du moins, la raison pour laquelle on a observé qu'un plus grand nombre de répondants remplissaient en entier le formulaire en ligne (par comparaison avec le formulaire papier) au lieu de se décourager et d'abandonner en cours de route.

38. Remplir en ligne un questionnaire complexe (présenté sous forme de matrice et s'adressant aux ménages) demeure encore relativement exceptionnel et peut être considéré par certains comme une nouveauté. Il est difficile de savoir si les mêmes impressions concernant la facilité d'utilisation et la qualité des résultats vont persister dès lors qu'un plus grand nombre d'utilisateurs, peut-être moins expérimentés, sont incités à utiliser l'application en ligne pour les futurs recensements.

B. Degré élevé de satisfaction des utilisateurs

39. Il est absolument indispensable de veiller à ce que l'utilisation du questionnaire en ligne procure une grande satisfaction au répondant afin qu'il poursuive la session jusqu'à son achèvement et renvoie un questionnaire entièrement rempli. Cette satisfaction est déterminée par un certain nombre de facteurs, par exemple la facilité d'accès au questionnaire, qui dépend des exigences minimales imposées par le système pour en assurer la sécurité et de la complexité de l'ouverture d'une session. Parmi les autres facteurs figurent notamment l'absence de téléchargement exagérément long ou d'empreinte, l'utilisation de la configuration par défaut de la plupart des machines, la compatibilité avec les navigateurs courants, des aiguillages automatisés intelligents, des procédures de vérification souples qui permettent de continuer à remplir le questionnaire même si une réponse particulière n'est pas valable ou n'a pas été

fournie, et l'impression pour le répondant que le formulaire est relativement rapide et facile à remplir. Après le test de recensement de 2004, Statistique Canada a procédé à une enquête de suivi auprès d'un petit sous-ensemble de répondants par l'Internet afin d'apprécier leur degré de satisfaction. Une partie des résultats de cette enquête sont indiqués ci-après:

- a) Dans leur majorité (89 %), les répondants ont déclaré qu'ils avaient rempli à leur domicile le questionnaire utilisé pour le test de recensement;
- b) Soixante-dix-neuf pour cent des répondants ont déclaré qu'ils disposaient d'une connexion «haute vitesse»;
- c) Lorsqu'il leur était demandé la raison pour laquelle ils avaient choisi de remplir en ligne le questionnaire utilisé pour le test de recensement, 52 % des répondants ont déclaré que c'était plus facile, 30 % que c'était plus rapide, 18 % qu'ils préféraient cette formule et 16 % qu'ils n'avaient pas à renvoyer le questionnaire par la poste;
- d) Quatre-vingt-quinze pour cent des répondants ont estimé qu'ils étaient globalement satisfaits d'avoir utilisé l'application Internet pour le recensement;
- e) Quatre-vingt-huit pour cent des répondants, dont plus de la moitié avait utilisé la version abrégée, ont jugé acceptable le temps nécessaire pour remplir le questionnaire utilisé pour le test du recensement. Ceux qui ont indiqué ne pas disposer d'une connexion «haute vitesse» avaient davantage tendance à déclarer qu'il fallait trop de temps pour remplir le questionnaire;
- f) Quatre-vingt-dix-huit pour cent des répondants ont déclaré qu'ils rempliraient le questionnaire de recensement en ligne en 2006;
- g) Cinquante-sept pour cent des répondants ne craignaient pas du tout que l'utilisation de l'Internet entraîne un risque accru pour la sécurité et le caractère confidentiel des données fournies en réponse au questionnaire utilisé pour le test de recensement, alors que 34 % s'en inquiétaient et que 8 % le craignaient beaucoup;
- h) Lorsqu'il leur a été demandé d'évaluer la sécurité de la transmission des données personnelles par l'Internet, 73 % des personnes interrogées ont déclaré que cette transmission était sûre, 13 % qu'elle était très sûre et 10 % qu'elle ne l'était pas du tout.

40. Trois enquêtes de suivi ont été réalisées après le recensement de 2006. Les résultats en seront communiqués pour le mois de mai 2007.

C. Possibilités sur le long terme

41. Si l'on pose un certain nombre d'hypothèses de base, par exemple le fait que la plupart des citoyens vont continuer d'accepter et d'utiliser l'Internet, l'option en ligne offre sur le long terme des perspectives d'économie, d'amélioration de la qualité et d'allègement de la charge de travail imposée aux répondants.

42. Dans les zones géographiques où le taux de pénétration de l'Internet est très élevé, il peut s'avérer matériellement possible d'offrir par défaut l'option de réponse en ligne, en envoyant

simplement l'URL et un code d'accès, ce qui permettrait peut-être de réduire sensiblement les frais d'impression et de distribution du questionnaire papier. Au Canada, les économies pourraient être proportionnellement plus importantes, étant donné que Statistique Canada doit le plus souvent imprimer et distribuer deux versions, anglaise et française, du questionnaire. La difficulté consiste à remettre en temps voulu un questionnaire papier dans la langue de leur choix à ceux qui ne peuvent pas ou ne veulent pas utiliser l'application en ligne (et le font savoir en téléphonant au service d'assistance).

43. Dans le cadre du recensement de 2006, une étude a été entreprise afin de mettre à l'essai une méthode visant à promouvoir l'option Internet. Un modèle a été élaboré en vue de déterminer les zones qui, a priori, comprennent un nombre important de logements correspondant à des ménages qui utiliseront probablement l'option en ligne pour le recensement. Les ménages considérés dans cette étude, qui a reçu le nom de «Stratégie de pression», n'ont reçu qu'une lettre au lieu du questionnaire papier. Il leur a été demandé de remplir le questionnaire en ligne. La lettre indiquait également un numéro 1-800 que les répondants pouvaient appeler pour s'informer au sujet de l'étude ou demander un questionnaire papier. Un échantillon préliminaire de 40 000 ménages résidant dans les zones dans lesquelles la lettre a été envoyée a été constitué aux fins de l'étude. Cet échantillon a été divisé au hasard en deux groupes de 20 000 ménages chacun afin de créer un groupe témoin. Les ménages témoins ont reçu un questionnaire papier et ceux de l'échantillon faisant l'objet de la Stratégie de pression ont reçu la lettre.

44. Les deux problèmes majeurs au cours de cette étude ont été l'incapacité d'identifier directement les ménages disposant d'une connexion Internet et la recherche des moyens d'alléger au maximum les opérations de collecte des données. À cet égard, l'échantillon n'était pas concentré mais plutôt réparti sur l'ensemble des zones cibles au Canada. Un modèle a été construit à l'aide des données provenant du test de recensement de 2004 et du recensement de 2001 afin de déterminer les zones comprenant un nombre non négligeable de logements dans le cas desquels le répondant allait probablement opter pour l'option en ligne. Le modèle était établi à partir des variables qui, sur le plan statistique, correspondait aux plus grandes différences entre l'option Internet et les autres modes de réponse. Par exemple, les variables telles que l'âge, le niveau d'instruction, l'activité sur le lieu de travail, la catégorie d'élèves ou d'étudiants, le nombre d'enfants à domicile, le revenu et les ménages multifamiliaux figuraient parmi les plus importantes. Ces variables déterminaient un «profil» Internet des personnes et des ménages. Ce profil était ensuite mis en regard des données de recensement de 2001 afin de donner une note à chaque individu puis à chaque ménage. L'analyse de la concentration des ménages associés aux notes les plus élevées a mis en évidence les zones dans lesquelles se trouvait la plus forte proportion de ménages qui allaient probablement opter pour la réponse en ligne. Enfin, les calculs réalisés pour 2001 ont été adaptés aux zones de 2006 à l'aide des fichiers de conversion géographique.

45. Cette méthode a été très efficace puisque le taux de réponse par l'Internet dans l'échantillon faisant l'objet de la Stratégie de pression a été 2,6 fois plus élevé que dans le cas du groupe témoin et 3,4 fois plus élevé que dans celui de la population en général. Dans la perspective du recensement de 2011, il est prévu de porter de 73 à 90 % la proportion des ménages appelés à recevoir une lettre et d'appliquer plus largement la Stratégie de pression.

46. Il est ressorti du recensement que le taux de réponse par question était beaucoup plus élevé et celui des rejets au contrôle très faible pour les questionnaires Internet. Un taux d'utilisation plus élevé entraînerait par la même occasion une diminution des taux de suivi et d'imputation, et de ce fait un abaissement des coûts et une amélioration de la qualité. L'application en ligne offre aussi des perspectives d'élargissement de l'éventail des choix offerts pour certaines réponses, par exemple dans les catégories relatives à l'emploi ou à l'instruction (profession, branche d'activité, niveau d'instruction), ce qui réduit le travail de codage manuel.

47. Eu égard à la diversité toujours plus grande de la population canadienne, l'Internet permet de présenter les questionnaires de recensement dans plusieurs langues moyennant un modeste surcoût. Même en cas d'utilisation du questionnaire papier par le répondant, l'Internet peut être utilisé pour la traduction des questions en plusieurs langues. Dans l'ensemble, ces fonctionnalités pourraient permettre d'accroître le nombre de réponses et de réduire les coûts de suivi. De même, l'application en ligne offre, pour les personnes qui souffrent d'un handicap, par exemple les aveugles, de véritables solutions de rechange sous forme d'autorecensement, au lieu de devoir faire appel aux services d'un agent recenseur.

VII. CONCLUSION

48. La mise en place par Statistique Canada d'une option de réponse en ligne pour le recensement de 2006 a donné de bons résultats. La réalisation en 2004 d'un vaste test de recensement en conditions réelles a fourni l'occasion de mettre à l'épreuve un concept novateur avec le concours des répondants dans le cadre d'un environnement intégré de collecte et de traitement des données. Un certain nombre de facteurs ont favorisé l'introduction d'une application en ligne pour le recensement de 2006, par exemple les taux de connectivité relativement élevés et la facilité de navigation sur le Web acquise par une majorité de la population canadienne. La méthode de collecte des données de recensement qui a consisté à envoyer les questionnaires par la poste ou à les déposer au domicile pour que les répondants procèdent à un autorecensement a également facilité l'introduction de l'application en ligne. La conception du traitement intégré qui a été mis au point a permis de faire converger tous les modes de collecte vers les mêmes systèmes et processus de vérification, de suivi et d'imputation en aval. Grâce à la décision de faire appel à des sous-traitants, un vendeur compétent, expérimenté et bien placé a fourni une application solide, intégrée et conviviale. L'expérience acquise en interne des enquêtes en ligne a permis de procéder à des essais ainsi que de spécifier des critères utiles et à satisfaire en priorité. La mise en place au bon moment de la voie de communication protégée par le Gouvernement canadien et l'attitude conciliante dont il a fait preuve pour l'adapter aux besoins de Statistique Canada a beaucoup contribué à l'établissement d'une structure sécurisée. Une solide structure de gouvernance et de gestion des projets a été créée afin que les modifications apportées au système et les coûts soient étroitement surveillés et strictement contrôlés. De la sorte, Statistique Canada s'est trouvé en bonne position pour répondre aux attentes des Canadiens qui souhaitaient disposer d'une application en ligne sécurisée, efficace et conviviale en 2006. Il va prendre appui sur cette expérience constructive pour le recensement de 2011.
