

Statistique Canada  
Division de l'agriculture

## **DOCUMENT DE TRAVAIL #34**

### **Utilisation de la technologie d'imagerie documentaire dans le recensement canadien de l'agriculture de 1996**

**Mel Jones et Ivan Green  
Statistique Canada**

Février 1998

N° 21-601-MIF98034 au catalogue

L'analyse et l'interprétation des données sont la responsabilité de l'auteur et non celle de Statistique Canada.

© Ministre de l'Industrie, Statistique Canada, 1998. Tous droits réservés. Il est interdit de reproduire ou de transmettre le contenu du présent document, sous quelque forme ou par quelque moyen que ce soit, enregistrement sur support magnétique, reproduction électronique, mécanique, photographique, ou autre, ou de l'emmagasiner dans un système de recouvrement.



# Utilisation de la technologie d'imagerie documentaire dans le recensement canadien de l'agriculture de 1996

Mel Jones et Ivan Green  
Statistique Canada  
le 3 février 1998

## Sommaire

Le traitement des données du recensement canadien de l'agriculture dépend grandement de l'accès à ses questionnaires. On a besoin de ces documents à bien des stades de la collecte et du traitement pour que des produits de qualité puissent être livrés à la clientèle agricole. L'accès physique aux documents s'est révélé difficile par le passé sur le plan de la remise des questionnaires en temps utile, de l'accès simultané, de la gestion, du repérage, de l'entreposage et des coûts. Dès 1984, les responsables du recensement ont reconnu que la technologie de l'imagerie documentaire pouvait les aider à améliorer la qualité et les délais de livraison des données. Il reste que, dans les années 1980, on se rendait compte d'emblée que la technologie du moment n'était pas assez mûre ni abordable pour le recensement. En 1996, l'imagerie documentaire est devenue une fructueuse réalité pour le recensement canadien de l'agriculture. Dans ce document, nous décrivons les expériences, les enjeux et les attentes des intervenants très divers qui ont été associés à l'implantation de cette technologie.

### *Introduction*

L'implantation de la technologie d'imagerie documentaire aux fins du recensement canadien de l'agriculture a comporté des frais et des risques de taille. Dès les premiers stades du projet, on a amplement discuté des méthodes classiques et des technologies rivales. Dans une première section, nous exposerons les questions débattues et dégagerons les orientations retenues. Nous mettrons en parallèle les inconvénients des méthodes manuelles du passé et les avantages attendus de cette technologie.

Nombreuses ont été les difficultés liées à l'implantation d'un système d'imagerie documentaire. Dans une deuxième section, nous évoquerons plusieurs des défis relevés et dirons comment les grands problèmes techniques qui se posaient ont été cernés et résolus. Suivront des observations au sujet des raisons, évidentes ou non, pour lesquelles ce projet d'implantation a été couronné de succès.

À la troisième section, nous parlerons de l'optique de la direction dans la décision prise de recourir à l'imagerie, ainsi que des possibilités de maintien et d'évolution de cette technologie à Statistique Canada.

### *Section I : Pourquoi recourir à l'imagerie?*

Selon la politique adoptée par l'organisme, les dates de diffusion des publications sont communiquées d'avance au public. Une fois déclarées, ces dates ne changent plus, aussi la production statistique est-elle assujettie à des échéances très fermes en fonction des dates de diffusion.

Les contraintes de temps et de budget, jointes à la nécessité de maintenir, voire d'améliorer la qualité, ont suscité un mouvement d'optimisation de l'efficacité. C'est pourquoi on mise davantage sur l'automatisation et l'innovation. Sont principalement visées ici les méthodes manuelles coûteuses.

Pour le recensement canadien de l'agriculture, on s'est surtout attaché a) à la saisie des données et b) à la remise des questionnaires tout au long des activités de production. On a réalisé des études de faisabilité de l'automatisation pour l'un et l'autre de ces aspects.

En ce qui concerne la saisie des données, on a étudié la technologie de reconnaissance intelligente de caractères (RIC). À l'époque, on s'interrogeait sur le mûrissement de la technologie RIC aux fins du recensement, et on manquait de temps pour vérifier les réactions des recensés. Si on avait utilisé cette

technologie, on aurait nettement alourdi le fardeau de réponse, car il aurait fallu passer de réponses numériques librement consignées par les recensés à des réponses numériques assujetties à des contraintes et adaptées à la technologie RIC. Comme plus de 90 % des renseignements sollicités dans le cadre du recensement sont numériques, on aurait sans doute eu l'obligation de repenser le questionnaire de fond en comble. Ajoutons que, après l'étude de faisabilité, on s'est posé des questions sur l'efficacité et la rentabilité des logiciels de reconnaissance de caractères.

La recherche sur l'imagerie documentaire et la technologie RIC a créé une situation très intéressante et riche en défis. Pour emprunter des termes de politique canadienne, disons que les membres de l'équipe de gestion du projet se sont divisés en trois camps, ceux des conservateurs, des progressistes-conservateurs et des progressistes. Les conservateurs jugeaient que l'implantation en soi d'un système d'imagerie était trop risquée pour les raisons suivantes :

- Le recensement allait être tenu dans deux ans et, dans une telle entreprise quinquennale, il était trop hasardeux de se mettre à essayer et à implanter un tel système en deux ans.
- Selon les prévisions, les coûts variaient amplement d'abordables à franchement inabordables.
- L'organisme ne connaissait rien de l'imagerie documentaire ni de la technologie RIC, aussi était-ce user de prudence que d'éprouver ces techniques dans des enquêtes plus modestes et plus fréquentes.
- Le personnel dépendrait entièrement des machines. Si le système tombait en panne, le personnel n'avait plus rien à faire.
- Le projet offrait de nombreux autres défis, dont les suivants :
  - On avait ajouté une nouvelle enquête sur l'état d'avancement des semailles. C'était là une enquête de suivi ITAO qui s'adressait à plus de 40 % des recensés.
  - Il fallait entièrement repenser les méthodes et passer d'une exploitation en ordinateur central à l'utilisation d'ordinateurs intermédiaires.
  - Il fallait un système et une méthodologie nouveaux pour le Registre des fermes.

Les conservateurs voyaient bien l'intérêt et les avantages de l'imagerie, mais en jugeaient les risques

trop grands pour le recensement de l'agriculture de 1996. La méthode manuelle du passé (sans imagerie) était éprouvée et la plus sûre.

Aux yeux des progressistes, on économiserait le plus en implantant conjointement un système d'imagerie et la technologie RIC. Sans cette technologie, le système d'imagerie ne serait pas rentable. Les tenants de ces vues étaient persuadés non seulement qu'un système d'imagerie était réalisable, mais aussi que la technologie RIC avait alors suffisamment mûri pour les besoins du recensement.

Le chargé de projet avait d'abord proposé d'implanter le système d'imagerie sans RIC. Les progressistes-conservateurs l'avaient appuyé. Ils jugeaient inacceptables les risques d'une implantation RIC dans les délais à respecter. Ils pensaient par ailleurs qu'un système d'imagerie était faisable sans trop de risques. On leur avait parlé d'une foule de projets d'imagerie qui avaient tourné à la catastrophe, mais ils avaient aussi pu observer assez d'applications qui fonctionnaient ailleurs pour acquérir la conviction que le jeu en valait la chandelle.

Dans ce débat, aucun camp n'avait tort ou raison, chacun ayant de bonnes raisons de prendre la position qu'il avait adoptée. Une fois la décision prise d'implanter un système d'imagerie sans technologie RIC, tous les membres de l'équipe de projet se sont totalement rangés à cet avis.

Le but était maintenant de communiquer des images électroniques des questionnaires aux ordinateurs de bureau des vérificateurs et des valideurs.

Si on avait exploité un système manuel de remise de questionnaires pour 1996, celui-ci aurait accusé les insuffisances de ceux qui l'avaient précédé :

- **Livraison en temps utile.** Il faudrait compter des heures, voire des jours pour la remise des questionnaires au personnel de production.
- **Accès simultané.** C'est là une autre question de livraison en temps utile. Souvent, les valideurs et les vérificateurs doivent avoir simultanément accès aux questionnaires de grandes exploitations agricoles. Ce sont ces dernières qui tiennent la plus grande place dans les estimations de nombreuses variables du recensement. La livraison serait successive, car on ne pourrait tirer des copies des questionnaires, activité refusée du fait de ses risques pour la sécurité. Pour contrôler les questionnaires, on les archivait et les désarchivait à chaque opération ou à chaque demande.

- **Repérage.** Pour contrôler le cheminement individuel des documents, il faudrait se doter d'un système de repérage automatisé. À chaque stade de la production, on aurait besoin d'accuser réception manuellement après un passage en lecteur de code à barres. Un système de ce genre imposerait un cheminement successif des questionnaires (et des données) dans le circuit.
- **Archivage complexe.** On aurait aussi besoin de tout un système de classement par chemises, boîtes et étagères pour un archivage et un désarchivage à grand volume.
- **Coûts élevés.** Avec les préposés, les étagères, les locaux d'archivage, les mesures de sécurité et la mise au point d'un système de distribution et de repérage de questionnaires, on se retrouverait avec un système manuel coûteux et relativement lent.
- **Frustration.** Les valideurs devaient dresser sur papier la liste des questionnaires dont ils avaient besoin et la présenter à la bibliothèque. En temps normal, on répondrait à leurs demandes en quelques heures, mais beaucoup de questionnaires manqueraient à cause de l'accès simultané, d'erreurs de classement, etc. Il faudrait produire plusieurs demandes avant que tous les documents désirés ne soient réunis.
- **Erreurs de classement.** Des erreurs de classement se produisent lorsqu'on archive et désarchive. Il faudrait interrompre périodiquement les opérations en bibliothèque pour effectuer une vérification systématique et exhaustive.

Par intuition, on pouvait voir qu'un système d'imagerie de questionnaires permettait de remédier aux insuffisances d'un système manuel :

- **Livraison en temps utile.** Il faudrait non plus des heures ou des jours mais quelques secondes pour communiquer des images de questionnaires au personnel de production.
- **Accès simultané.** Tous les utilisateurs pourraient examiner simultanément un même document, d'où des possibilités de traitement en *parallèle*.
- **Repérage.** Tout au long du traitement, on n'avait plus besoin de système manuel de repérage de questionnaires. Les gens n'avaient pas non plus à inscrire les entrées et les sorties de questionnaires.
- **Archivage complexe.** On aurait encore besoin d'un système temporaire d'archivage, mais il serait relativement plus simple.
- **Coûts élevés.** Comme on n'avait jamais fait d'imagerie à grande échelle à Statistique Canada, on n'avait pas de point de comparaison pour le

recensement de l'agriculture, aussi ne savait-on au juste si un système d'imagerie coûtait plus cher ou moins cher qu'un système manuel de remise de questionnaires.

La direction a jugé que les avantages possibles méritaient un examen plus attentif.

Outre les avantages énumérés, elle prévoyait les retombées suivantes :

- on n'aurait pas besoin de microfilmer les documents après traitement.
- on réduirait les opérations manuelles, notamment celles des préposés à la bibliothèque.
- grâce à un accès immédiat aux documents, le personnel gagnerait en efficience, car il pourrait constater les problèmes et les résoudre immédiatement; on n'aurait plus besoin de dresser la liste des problèmes et de trouver autre chose à faire jusqu'à ce que les documents arrivent.
- les commis et les valideurs auraient tous moins de papier à manipuler; la paperasserie serait en grande partie éliminée par la communication d'images électroniques.
- on n'aurait plus la frustration que causent l'attente de questionnaires et les documents en retard ou manquants.
- enrichissement des tâches : le personnel jugerait désormais satisfaisantes et stimulantes les tâches fastidieuses du passé.
- en ayant un accès immédiat aux questionnaires problèmes, les préposés compteraient plus sur l'information reçue que sur leur intuition; pour ne pas avoir à attendre de documents sur papier, ils auraient parfois tendance à deviner les problèmes et les solutions.
- grâce à un accès immédiat, le personnel pourrait considérer plus de problèmes et ainsi améliorer la qualité des estimations, notamment à de bas niveaux de description géographique.

## Section II : Questions d'imagerie

Lorsqu'on s'est mis à étudier l'imagerie documentaire, on a voulu répondre aux questions suivantes :

- Qu'est-ce que l'imagerie documentaire?
- Quels éléments d'imagerie permettront-ils de combler nos besoins?
- Cette imagerie fonctionnera-t-elle dans notre cadre technique?

- À quoi pourraient se monter les coûts en gros?
- Quelles questions doit-on étudier si on veut que l'implantation soit réussie?

Les représentants du secteur du recensement ont conféré en première étape avec le secteur privé. Ils ont d'abord communiqué à l'industrie les critères, les volumes et les spécimens du questionnaire de 1991 qu'ils jugeaient représentatifs du questionnaire de 1996 (qui n'avait pas encore été élaboré). Les représentants de l'industrie sont revenus avec des prévisions de coûts et des solutions proposées.

Selon leurs estimations, il fallait prévoir de 500 000 à presque 3 000 000 dollars canadiens. Le premier chiffre était acceptable et le second non.

Les propositions de solution prêtaient quelque peu à confusion. On y abordait des problèmes qui n'avaient nettement rien à voir avec le recensement et en oubliait d'autres qui risquaient de se poser. Il reste que, malgré un peu de confusion, les représentants du secteur du recensement ont commencé à se faire une meilleure idée des enjeux.

Voici notamment des questions qui semblaient se poser (et dont la solution serait fort coûteuse) pour le recensement :

- indexation;
- contrôle de la qualité;
- débit des lecteurs automatiques sur papier ministre Vista 70M (questionnaire du recensement);
- modifications à prévoir à la disposition des questionnaires pour les adapter au **systeme d'imagerie**;
- qualité des images.

Pour pouvoir dresser une liste suffisante de questions, on a assisté à des séminaires sur l'imagerie documentaire dans le cadre des salons spécialisés de l'Association for Information and Image Management (AIIM) et rendu visite à des chercheurs en imagerie du United States Bureau of the Census (USBC).

Les représentants de l'USBC ont mis en lumière la question primordiale de l'**intégrité des documents**. Les questionnaires de cet organisme se présentent sous forme de livret (plusieurs pages). Les lecteurs traitent un feuillet à la fois de sorte que, pour capter un document de plusieurs pages, on peut être obligé de démembrer un questionnaire en autant de pages. Ce qui inquiétait avant tout les chercheurs de l'USBC, c'était le risque que des pages de tels documents s'égarerent ou ne soient plus dans le bon

ordre avant ou après la lecture, c'est-à-dire le risque de perte d'intégrité des documents. À l'époque, l'USBC étudiait des dispositifs tourne-pages (plutôt que coupe-pages). Comme le questionnaire du recensement de l'agriculture comprend plusieurs pages, on a vu que le problème se posait aussi à Statistique Canada. Ces questionnaires étant relativement peu nombreux (280 000), les dispositifs tourne-pages ne paraissaient pas une solution rentable.

Une des premières solutions envisagées a été d'apposer un code unique à chaque page d'un questionnaire. On s'est entretenu à ce sujet avec des spécialistes internes et des représentants de grandes imprimeries. La solution était certes réalisable, mais son rapport efficacité-coût était inacceptable, puisqu'on devait prévoir près de 400 000 dollars canadiens, ajoutant ainsi 1,43 \$ environ au coût d'impression de chaque questionnaire, ce qui ferait plus que tripler le budget d'impression. Avec cette solution, il y avait toujours des risques d'erreur de mise en ordre. Comme autre solution, on a songé à faire imprimer en relief un numéro d'ordre au bas de chaque page par lecteur automatique. C'était là la bonne solution pour les questionnaires égarés **après** lecture, mais non avant lecture. Il fallait donc pousser l'étude de la question.

On a continué à examiner la question de l'imagerie documentaire. Les salons spécialisés nous ont éclairés sur d'autres questions :

- taille et qualité des moniteurs;
- cartes vidéo et fonctionnalité;
- capacité réseau;
- capacité de stockage;
- fonctionnalité des lecteurs;
- soutien technique;
- intégration aux applications internes.

On a fini par dresser une liste relativement exhaustive, mais on ne savait encore au juste dans quelle mesure ses éléments s'appliquaient tous au recensement de l'agriculture. On craignait que l'application du recensement comporte des questions particulières d'imagerie à régler. Ainsi, la plupart des lecteurs sont conçus pour traiter des feuillets uniques format lettre de force 30. Pour le recensement, on utilise des livrets de 16 pages de papier ministre Vista 70M. La direction a accepté que l'on réalise un petit projet pilote (au coût de 20 000 dollars canadiens) pour clarifier les questions relatives au recensement.

On a demandé à une entreprise locale d'implanter à contrat un système d'imagerie de base sur place. Elle a travaillé en étroite collaboration avec les responsables des systèmes internes et les spécialistes des méthodes. L'exercice visait :

- à étudier les méthodes et les procédures d'exécution nécessaires à l'implantation d'un système d'imagerie entièrement fonctionnel pour ce qui est de la réception, de la préparation, de la lecture, du remontage, de l'archivage, du désarchivage et du contrôle de la qualité des questionnaires;
- à apprendre à connaître le matériel et les logiciels types d'imagerie de l'industrie;
- à se renseigner sur le rendement de cette technologie dans le cas des documents du recensement.

Le système pilote se composait des éléments suivants :

- deux lecteurs (Fujitsu et Bell + Howell);
- un serveur Hewlett-Packard sur UNIX et LAN-Manager 2.1;
- quatre PC dotés de cartes vidéo de 2 Mo;
- moniteurs couleur et échelle de gris de 17 et 21 po;
- TCP/IP;
- sous-réseau Ethernet 10 Base T UTP;
- visualisateurs d'imagerie Watermark TIF;
- échantillon de 2 000 questionnaires de 16 pages papier ministre du recensement de 1991.

Le projet pilote a duré trois jours et donné les résultats suivants :

- La technologie a été «politiquement» mieux acceptée. L'observation des questionnaires en lecture, de la qualité des images et de la fonctionnalité des visualisateurs a accru la confiance des gens et atténué les réactions de scepticisme.
- Comme Statistique Canada s'était conformé aux normes d'interface, il n'y a pas eu de surprises tenant à l'infrastructure technique.
- Le débit observé des lecteurs était considérablement moindre que prévu, mais impressionnait quand même.
- Dans cette expérience faite dans le secteur du recensement, l'équivalence de modèles de lecteurs bien entretenus venant de fabricants concurrents n'est pas synonyme d'équivalence de rendement. Un des lecteurs a été d'un fonctionnement sûr avec relativement peu de captation en double, d'images obliques et de blocage. L'autre a été d'un rendement fort

médiocre. Comme il détruisait littéralement les questionnaires, on l'a rapidement retiré du projet pilote.

- La fonctionnalité des lecteurs était très coûteuse (lecture de codes à barres, indexation, comptage, etc.).
- On a pu établir le nombre de lecteurs de ce type et le nombre de préposés dont on avait besoin pour les quantités de questionnaires de recensement à traiter.
- La qualité des images était acceptable (200 ppp).
- On a déterminé la taille moyenne des images (produites par les lecteurs du projet pilote); avec ces indications, on pourrait mieux simuler le trafic réseau et établir les besoins de stockage d'images.
- La force et la couleur du papier des questionnaires (papier Vista opaque 70M) étaient acceptables; il n'y avait pas de «lavage» de caractères d'une extrémité à l'autre d'un feuillet.
- La disposition du questionnaire de 1991 ne se prêtait pas à l'imagerie, mais les corrections à apporter seraient légères et n'auraient aucune incidence sur les recensés ni sur le budget d'impression en 1996. On n'avait pas à vérifier la réaction des recensés à un questionnaire adapté à l'imagerie.
- La fonctionnalité, l'aspect et l'impression qu'offre un visualisateur Watermark seraient conformes aux exigences du recensement, ce qui permettrait d'estimer grossièrement les coûts afférents et d'obtenir des valeurs de référence à des fins de spécification technique.
- Le problème d'intégrité des documents ne se posait pas. On avait craint que, si des documents coupés tombent, des pages soient égarées ou ne soient plus dans le bon ordre, d'où une perte d'intégrité. Des essais ont démontré que les liasses de questionnaires coupés ne se défaisaient généralement pas si elles tombaient à cause de l'électricité statique et des plis des pages. Le risque que des pages se perdent ou s'égarer n'en subsistait pas moins. On a pris des mesures pour l'atténuer encore plus. On a notamment gardé les liasses petites (30 documents au plus), les a rangées dans des enveloppes en réduisant au minimum le temps passé hors pli, diminué la distance entre le massicot et le lecteur, fait le remontage des documents avec des caractères manuscrits et des couleurs d'encre caractéristiques et, comme dernière précaution, consulté les images ou les données de la base de données structurées en vue du remontage des documents qui étaient tombés.

- Des moniteurs couleur 17 po moins chers avec leurs cartes vidéo étaient acceptables (résolution 1 024x768) pour au moins la moitié des activités de production du recensement. L'image était souvent d'une qualité *supérieure* à celle de l'original, parce que les images pouvaient être grossies. Avec les moniteurs à échelle de gris, l'affichage était plus clair, mais pas outre mesure. On jugeait en outre plus facile d'affecter les moniteurs couleur à d'autres activités après la production.

Les résultats du projet pilote ont frayé la voie à l'implantation du système. Voici ce qu'il restait à faire :

- élaboration d'une stratégie de contrôle de qualité;
- élaboration d'une stratégie d'indexation;
- établissement des besoins réseau;
- établissement des besoins en intégration;
- établissement des besoins en lecteurs;
- établissement des besoins en visualisateurs;
- choix et acquisition de moniteurs et de cartes vidéo;
- acquisition d'une technologie de stockage;
- conception de systèmes d'exécution et de contrôle de lecture;
- établissement d'une demande de propositions (DDP);
- implantation.

Dans tout le reste du projet, les contributions respectivement apportées par les représentants de la Division des services aux utilisateurs de l'informatique (DSUI), du Groupe de soutien aux ordinateurs de gamme moyenne, de la Division des méthodes d'enquêtes-entreprises et de la Division du développement des systèmes de Statistique Canada ont été primordiales.

**Largeur de bande :** En ce qui concerne les charges de production en imagerie, un grand sujet d'inquiétude était le réseau. Quelques années avant le projet, l'organisme s'était recâblé en 10 Base T UTP. Le nouveau câblage s'était révélé vulnérable lorsque de grands fichiers étaient transmis. On devait s'assurer qu'il pouvait soutenir les activités simultanées de 100 utilisateurs finals et des opérations de lecture à grand volume. L'absence de point de comparaison dans l'organisme pour les charges de traitement d'images posait un problème. Heureusement, on pouvait toujours consulter les principaux responsables de l'exécution du recensement précédent et s'enquérir des quantités de questionnaires qui étaient quotidiennement passées

par le système manuel. On prévoyait que les volumes d'images dépasseraient ceux des documents sur papier. Pour vérifier, on a triplé les quantités traitées comme trafic utilisateur. On a ajouté cette quantité au volume prévu de trafic lecteur (nombre total de questionnaires divisé par le nombre total de jours disponibles de lecture) pour représenter la charge quotidienne totale du réseau en imagerie. On a divisé ce total par 6,5 heures (nombre moyen d'heures d'activité dans une journée de 7,5 heures) pour obtenir la charge moyenne en cours de production. À l'aide de cette formule et des données sur la taille des images fournies par le projet pilote, on a mené divers essais réseau.

On a sollicité l'aide des spécialistes du Centre de services en communication (CSC). Ceux-ci ont indiqué comment s'y prendre efficacement pour réaliser des essais qui donneront les résultats les plus utiles. Pour représenter l'activité de 100 utilisateurs, il ne suffisait pas de simuler la demande en mégaoctets-seconde, c'est-à-dire le fonctionnement d'un poste de travail en communication d'imagerie dans les deux sens avec un serveur dans le réseau. Il importait aussi de simuler la répartition de la demande. Si un seul poste faisait venir des images et un autre en envoyait, des collisions étaient improbables en réseau. Il fallait donc se servir d'un grand nombre de postes pour que les essais soient significatifs.

Avant de procéder à ces essais, le CSC a configuré un sous-réseau isolé. Il a éliminé le bruit pour que les charges mesurées soient attribuables à la transmission d'images. On a fait marcher 20 postes de travail avec chacun un programme qui téléchargeait un fichier en provenance ou à destination d'un serveur dans un laps de temps donné. Le CSC a pris des analyseurs de réseau appelés «renifleurs» pour observer le trafic.

Il a recommandé l'acquisition d'un autocommutateur à routage Ethernet d'une capacité de 100 Mo/s pour les charges prévues en production. Des représentants du Groupe de soutien technique ont installé l'autocommutateur et isolé les utilisateurs du recensement du reste de l'organisme. Le but était double : (1) protéger le sous-réseau d'imagerie contre toute dégradation du rendement imputable à des variations imprévues du trafic de réseaux locaux extérieurs et (2) protéger ces réseaux contre les charges d'images du secteur du recensement.

On a réparti les utilisateurs du recensement en trois groupes de travail dans l'autocommutateur Ethernet avec une liaison de 10 Mo/s. On a relié le serveur à

l'autocommutateur par un câble optique de 100 Mo/s. Le secteur du recensement a payé une carte FDDI (Interface optoélectronique de données à mode réparti) pour l'interface serveur-réseau. Le groupe de soutien aux ordinateurs de gamme moyenne a étudié les solutions possibles, présenté une recommandation et, après consultation, fait l'acquisition et l'installation de la carte FDDI.

Après l'implantation du *sous-réseau d'imagerie*, le CSC a procédé à de nouveaux essais. Il a ainsi pu vérifier que les besoins avaient bel et bien été comblés.

**Moniteurs :** Comme le système d'imagerie est au service de l'utilisateur final et qu'il n'est souvent connu que par ce qui paraît à l'écran, le moniteur en est sans doute l'élément primordial, et sûrement un des plus coûteux. Pendant le projet pilote, on n'a relevé aucune différence importante entre le moniteur couleur de 17 po et le moniteur échelle de gris de 19 po pour la qualité de perception des images du recensement. Comme les moniteurs devaient être utilisés dans plusieurs applications d'imagerie, on a décidé de faire l'acquisition de moniteurs couleur malgré leur prix supérieur. Il restait à établir a) la (les) taille(s) d'écran convenant aux besoins du recensement et b) les exigences relatives aux cartes vidéo.

À certains stades des activités de production, on doit prévoir l'observation simultanée de deux pages format ministre ou plus. On ignorait si un moniteur de 17 po avec une carte vidéo bien configurée suffirait ou non. L'écart de prix est très appréciable entre les écrans de 17 et de 21 po. Il s'agissait de voir si on avait effectivement besoin de grands moniteurs chers.

Le personnel du recensement disposait alors de plusieurs moniteurs de 17 po pouvant servir à une évaluation en imagerie, mais il n'y avait pas de moniteurs plus grands. On a pris contact avec des représentants de l'entreprise privée pour emprunter des moniteurs de 21 po.

Les membres du personnel de production du secteur du recensement qui s'occupaient des aspects plus techniques ont offert d'évaluer les moniteurs et les cartes vidéo. On a réparti par roulement les moniteurs grands et moyens entre les intéressés pendant trois semaines. On a ainsi pu établir que, malgré une *capacité virtuelle* d'écran et une résolution de 1 280x1 024, les moniteurs de 17 po ne suffisaient pas à tous les besoins du recensement. Les

cartes vidéo permettent des configurations virtuelles d'écran. Ainsi, l'utilisateur peut spécifier des aires d'affichage supérieures aux dimensions de l'aire d'affichage réelle de son moniteur. En maniant la souris, on peut faire apparaître les parties cachées de l'écran virtuel. On a jugé que cet écran était essentiel à certaines des activités sur moniteur de 21 po. Avec ces indications de recherche, on a pu arrêter des quantités de grands et de moyens écrans. Il a ensuite fallu passer aux considérations budgétaires.

Le prix, les dimensions et le rendement des moniteurs varient considérablement. On devait s'assurer que des moniteurs de 21 po pourraient être commodément disposés sur les bureaux, chose impossible dans certains cas. Aspect peut-être plus important encore, la qualité de l'affichage (netteté, contraste, luminosité et uniformité des images dans toute l'aire d'affichage) variait amplement pour des moniteurs d'une configuration et d'un prix comparables.

L'organisme maintient des offres à commandes (OAC) avec les fournisseurs en vue de faciliter ses acquisitions en technologie informatique. Sur la liste OAC, il y avait des appareils offerts par un certain nombre de fournisseurs. Aux fins du projet, on avait intérêt à exploiter cette liste si, parmi les moniteurs qui y figuraient, certains répondaient à nos besoins. Un représentant du groupe de soutien des postes de travail de l'organisme nous a fourni des appareils d'emprunt correspondant à tous les modèles de moniteurs 21 po de la liste. On a utilisé simultanément ces moniteurs à un même poste de travail et les mêmes images y ont été affichées. Le prix des appareils variait de 900 à 3 000 dollars canadiens. De prime abord, le moniteur le moins cher a bien paru, mais à y regarder de plus près, on a constaté que les couleurs se brouillaient et s'estompaient aux bords de l'écran. On a également observé un certain scintillement à une résolution de 1 024x1 280. Les moniteurs plus chers marchaient très bien, mais ils étaient d'un prix supérieur et d'une qualité inférieure par rapport à un des modèles déjà évalués par le personnel du recensement. On s'est arrangé avec le fournisseur pour qu'il ajoute ce produit à la liste OAC. (Les cartes vidéo désirées s'y trouvaient déjà.) Ces moniteurs ont été compris dans une commande de 120 Pentium destinés au personnel de production.

Pendant les activités de production, on a remarqué que les utilisateurs finals étaient nombreux à ne pas exploiter la capacité d'affichage virtuel des cartes vidéo simplement parce qu'ils ne s'y sentaient pas à l'aise. Ajoutons qu'une foule d'utilisateurs n'ont pas

utilisé les valeurs supérieures de résolution permises par les moniteurs et les cartes vidéo. Beaucoup ont choisi une résolution équivalente à celle de leurs anciens moniteurs SVGA de 14 po. Peut-être les petits caractères nets que l'on peut obtenir avec cette technologie sont-ils trop menus pour se lire facilement, surtout si les gens souffrent d'une déficience de la vue.

**Contrôle de la qualité :** Dans bien des systèmes d'imagerie, il y a des postes de travail et des préposés qui veillent sur la qualité des images communiquées au personnel de production. Les mesures de contrôle de la qualité (CQ) peuvent largement ajouter au coût de systèmes d'imagerie documentaire, car on doit souvent s'attendre à acquérir du matériel coûteux, à verser des salaires en plus et à subir des engorgements.

Les goulots d'étranglement sont particulièrement à craindre. La plupart des opérations en imagerie documentaire reposent sur une production d'images, et les opérations du recensement, sur une production de données. Ainsi, lorsqu'un vérificateur ou un analyste a accès aux données saisies dans la base de données structurées, il est prêt à commencer son travail et, si l'image du questionnaire correspondant n'a pas encore été communiquée, il travaillera sans cette image (auquel cas le système ne sert plus à rien) ou attendra qu'elle soit produite. Aux fins du recensement, il est impérieux d'obtenir les images **avant les données structurées ou en même temps**. Pour pouvoir examiner (en temps utile) les quantités quotidiennes d'images produites par les lecteurs sans retard de production, il aurait fallu disposer d'un système coûteux de contrôle de qualité.

Par des discussions avec la direction du recensement, les agents principaux et les spécialistes des méthodes et des systèmes, on a cherché des façons de contourner le problème. Tant qu'une page restait lisible, peu importait à l'utilisateur si le reste du document ne l'était pas. Le plus souvent, l'utilisateur final ne s'intéressait qu'à une partie de page de chaque document de 16 pages qu'il demandait. Un spécialiste des méthodes a examiné les données d'évaluation des lecteurs automatiques (statistiques sur les cas de captation en double et les cas d'obliquité, de déchirure et de pliure des images) et prévu le nombre de fois qu'un utilisateur final rencontrerait une image inacceptable et demanderait une nouvelle mise en lecture. Ce nombre était négligeable, car les images produites par les lecteurs étaient invariablement d'une grande qualité. C'est ainsi qu'on a décidé de contrôler la qualité après les

activités de production, le temps étant un facteur critique en production. La décision était d'autant plus facile à prendre que, pour demander une nouvelle lecture, il suffisait à l'utilisateur de **cliquer sur un bouton**.

Comme on a fait preuve d'imagination et s'est attaché aux besoins plutôt qu'aux solutions techniques classiques, on a largement pu réduire l'ampleur et le coût du contrôle de la qualité. On n'a pas eu besoin de postes de travail supplémentaires, et les préposés ont fait du contrôle de qualité pendant les **heures de relâche en production**.

**Indexation :** Comme dans d'autres types de bases de données, on emploie une ou plusieurs fonctions en relation avec les images pour une extraction en temps utile par l'utilisateur final. Dans certaines applications, on extrait automatiquement par lecture optique (reconnaissance optique de caractères ou ROC) les données d'identification des images avec lesquelles on dresse ensuite des index.

Comme les données d'identification des questionnaires étaient librement consignées par les recensés (elles n'étaient donc pas dactylographiées ni produites par imprimante), la ROC n'avait rien d'une solution applicable au recensement canadien de l'agriculture. On avait déjà écarté la RIC. Une autre possibilité était de se doter d'une application d'indexation interactive où le personnel devrait visualiser certaines pages de chaque image de questionnaire et saisir les données d'identification ensuite indexées (documents visualisés). Il y avait à cela un double inconvénient, puisque (1), avec des postes d'indexation, on se trouvait à accroître le coût du système d'imagerie et (2) on risquait de créer des goulots d'étranglement avant même les activités de production.

On pouvait heureusement obtenir les index nécessaires ailleurs et sans alourdir le budget. On avait tous les codes appropriés de renvoi à la base de données saisies des questionnaires. Il fallait une liaison entre le système d'imagerie et cette base d'information. On a demandé aux représentants de la Division du développement des systèmes de Statistique Canada d'étudier une solution.

On savait que les lecteurs pouvaient lire les codes à barres et les enregistrer automatiquement comme éléments d'indexation des images. Suivant les plans de saisie des données, des étiquettes de code à barres devaient être apposées aux questionnaires lors de l'introduction des données et saisies avec les données

des recensés. Ce serait là le code unique d'indexation versé dans la base de données.

En fonction de ce scénario, les spécialistes des systèmes ont proposé d'élaborer un programme permettant a) de consulter les nouvelles images de la base d'imagerie pour en extraire un code à barres et un code propre à cette base et b) de repérer, par le code à barres, l'enregistrement correspondant de la base de données structurées et y stocker le code interne de la base d'imagerie. L'utilisateur final aurait ainsi la possibilité d'accéder à un enregistrement de la base de données structurées par une des fonctions prédéfinies, puis de sélectionner l'image correspondante de la base d'imagerie à l'aide du code interne de cette dernière. C'était là une solution efficace et relativement peu compliquée.

**Forme du système :** Ce qui était très obscur dans la solution de l'imagerie documentaire était en train de se dissiper. Comme on n'avait besoin ni d'applications CQ, ni d'applications d'indexation, ni d'opérations de numérotation des questionnaires par gaufrage, on semblait se retrouver avec les seuls logiciels essentiels suivants :

- logiciel de gestion de lecture;
- logiciel de gestion de la base d'imagerie;
- programme de visualisation;
- logiciel de raccordement du système d'imagerie et de l'application du recensement (intégration).

Le matériel nécessaire était le suivant :

- lecteur(s);
- serveur;
- unité(s) de stockage;
- infrastructure réseau;
- postes de travail, moniteurs et cartes vidéo.

Quant à l'élément papier (questionnaire), on avait réglé la question de la couleur et de l'*adaptation au système d'imagerie*.

### **Logiciel de gestion de lecture automatique :**

On avait beaucoup appris au sujet de la fonctionnalité et du rendement des lecteurs par le projet pilote, les salons spécialisés et les discussions avec l'industrie, mais les évaluations avec le papier du recensement n'avaient pas visé un des chefs de file de l'industrie, la société KODAK. L'oubli semblait de taille, car KODAK avait autant de chances d'être choisie comme soumissionnaire DDP que les autres fabricants de lecteurs automatiques. On n'avait pas arrêté les détails essentiels (on savait seulement que le code à barres serait le code d'identification) du mode

de gestion des documents du recensement mis en lecture.

Comme les modèles KODAK étaient relativement hauts de gamme par rapport aux lecteurs du projet pilote, on aurait sans doute besoin de moins de préposés à la lecture automatique. Un représentant de KODAK et un gestionnaire du recensement chargé de l'opérationnalisation des fonctions de lecture ont eu un entretien. Ils ont organisé l'emprunt de l'ImageLink 500 qui serait utilisé dans les locaux du recensement. Les représentants du secteur du recensement, un spécialiste des méthodes, des spécialistes des systèmes, des membres du groupe de soutien en micro-informatique et les représentants de KODAK ont travaillé en étroite collaboration pendant un mois. Il s'agissait non seulement d'apprendre à connaître la fonctionnalité et le rendement des modèles KODAK, mais aussi de comparer ces appareils à ceux des autres fabricants qui avaient été employés dans le projet pilote. On sentait de plus en plus quelle technologie de lecture automatique répondrait le mieux aux besoins du recensement.

On a conçu les essais pour qu'ils portent sur tous les aspects imaginables de la lecture automatique des documents du recensement. La gestionnaire des opérations, son adjointe et un spécialiste des méthodes :

- se sont attachés au débit des lecteurs;
- ont étudié la qualité des images;
- ont jugé si une lumière blanche suffisait au balayage des deux couleurs du questionnaire;
- ont appris ce qu'il fallait faire pour assurer le fonctionnement et l'entretien des lecteurs.

Voici les données primordiales qui ont été tirées de cet exercice :

- (1) spécifications fonctionnelles des lecteurs à inclure dans la DDP;
- (2) chiffres d'analyse comparative (nombre de cas de captation en double, d'obliquité des images et de blocage);
- (3) tailles moyennes d'images KODAK acceptables (tailles qui s'accordaient avec celles des images produites dans le projet pilote);
- (4) précisions sur le plan des besoins en dotation;
- (5) constatation que le mécanisme de transmission KODAK travaillait mieux avec des demi-lots (15 documents de recensement) qu'avec des lots entiers (30);
- (6) constatation que les codes à barres doivent être très lisibles pour que le lecteur les reconnaisse. Ces codes ne sont pas tous d'une qualité égale. Sans cette évaluation pratique, on aurait peut-être

fait cette constatation essentielle à un moment où tout aurait été paralysé.

À la suite de cet exercice, les responsables de la lecture automatique se sont attaqués aux difficultés considérables de l'opérationnalisation des fonctions de lecture. Il fallait notamment concevoir des opérations de réception, de tri, de balayage et de stockage éventuel des documents. On a ainsi conçu et spécifié un système automatisé de contrôle du cheminement des documents en temps normal et en cas de demande de nouvelle lecture. Un spécialiste des systèmes de l'organisme a construit ce système en Visual Basic, a participé aux essais dont il a fait l'objet et l'a installé. En production, celui-ci a été très efficace et d'un fonctionnement très sûr.

Pendant l'évaluation KODAK, un gestionnaire principal des tâches de recensement s'est occupé de l'élaboration et de la diffusion de la DDP. Pour ce faire, il a eu recours à la gestionnaire des opérations, à son adjointe, à l'agent de liaison technique du recensement, à un membre du groupe de soutien aux ordinateurs de gamme moyenne et à un représentant de la Division du développement des systèmes. La demande de propositions devait comprendre les éléments suivants :

1. renseignements généraux sur le recensement;
2. renseignements sur l'infrastructure technique (technologie serveur, système d'exploitation, protocole réseau, configuration et système d'exploitation d'informatique de bureau, etc.);
3. critères de fonctionnalité et de rendement des lecteurs;
4. besoins de formation;
5. besoins d'intégration;
6. contraintes techniques relatives à la taille des images;
7. exigences fonctionnelles de visualisation;
8. besoins de soutien;
9. critères d'acceptation;
10. délais.

Pour lancer une DDP dans la fonction publique fédérale, il faut suivre un grand nombre de règles et de règlements. Quelque nécessaire qu'il soit, cet exercice demande d'ordinaire jusqu'à six mois entre la remise d'une DDP aux Travaux publics et l'officialisation du choix d'un soumissionnaire. Le temps devenait un facteur critique dans le projet, les activités de production devant débuter dans moins de huit mois. Nous croyons que le processus DDP aurait pu compromettre la réalisation du projet si le responsable n'était pas toujours resté en contact avec les Travaux publics et n'avait pas fait pression sur ce

ministère pour que la DDP du recensement demeure hautement prioritaire.

**Visualisateurs :** On se retrouvait devant l'alternative de faire l'acquisition de visualisateurs ou d'en construire sur mesure. Certains représentants de l'industrie pensaient qu'on économiserait considérablement en prenant l'option «visualisateurs sur mesure». D'autres étaient contre à cause des problèmes que posent les systèmes maison. On ignorait si le soumissionnaire retenu créerait ses propres visualisateurs ou en utiliserait qui existent déjà. D'une manière ou d'une autre, il fallait établir les spécifications fonctionnelles de visualisation et les inclure dans la DDP.

Cette détermination s'est révélée relativement simple. On a obtenu un visualisateur Watermark et l'a mis à la disposition de certains préposés au recensement. Après une démonstration et une période d'évaluation du logiciel, on a procédé à une «revue générale», occasion pour les intéressés d'indiquer simplement ce dont on avait besoin à leur avis et ce qu'il fallait modifier pour ensuite dégager des priorités. Les agents ont également précisé ce qu'ils jugeaient inutile. Les résultats de la consultation ont été documentés et intégrés à la DDP. On s'assurait ainsi que les besoins en visualisateurs seraient comblés par des produits standard ou une technologie sur mesure conçue par le sous-traitant.

**Besoins de stockage :** Cette détermination a été relativement simple. Après avoir examiné des échantillons d'images obtenues dans le projet pilote et dans l'évaluation des modèles KODAK, on a établi une taille moyenne d'image d'environ 600 K, valeur qu'on a multipliée par le nombre de questionnaires (en prévoyant une marge de sûreté de 25 %) pour arrêter l'ordre de grandeur du besoin global de stockage.

Il y avait toutefois là un certain risque, car on devait lancer la DDP de la technologie de stockage avant d'avoir choisi un soumissionnaire pour la DDP du système d'imagerie. La taille de l'image influait directement sur trois aspects primordiaux :

1. **Qualité des images :** Plus une image est grande, plus elle est de qualité. Il fallait veiller à ce que les images soient acceptables.
2. **Besoins de stockage :** Avec de grandes bases d'imagerie (4,8 millions d'images), une infime variation de la taille des images peut avoir une nette incidence sur les besoins de stockage. Il

fallait disposer d'une capacité de stockage suffisante avant les activités de production.

3. **Rendement** : Plus une image est grande, plus la charge est lourde pour le réseau, le serveur et le poste de travail. On avait fait tous les essais préalables de rendement en fonction d'une taille prévue des images. Si cette taille devait nettement augmenter, on aurait peut-être à en subir les conséquences.

**Rendement** : Un souci premier des gestionnaires de la production au recensement était celui de la vitesse de communication des images demandées aux ordinateurs de bureau. C'était peut-être là toute la différence entre la réussite et l'échec du système d'imagerie.

On avait pour défi de réduire les temps de communication au minimum. Le problème se posait lorsqu'il s'agissait d'une technologie à couches multiples qui transmettait les images et que d'autres activités se disputaient les capacités de calcul et d'acheminement réseau.

Voici les éléments de rendement que l'on devait examiner :

- Ordinateurs de bureau :
  - Vitesse de l'horloge;
  - Mémoire vive;
  - Antémémoire;
  - Espace de permutation;
  - Système d'exploitation (16 ou 32 bits);
- Moniteurs et cartes vidéo;
- Réseau :
  - Capacité (10 ou 100 Mo/s);
  - Protocole (TCP/IP, tailles des paquets, etc.);
  - Topologie (trafic isolé ou non);
- Images :
  - Taille des images;
  - Taille de la base d'imagerie;
  - Technologie de base de données;
- Serveur :
  - Mémoire vive;
  - Système d'exploitation;
  - Serveur réservé ou soutenant d'autres applications;
- Unités de stockage :
  - Antémémoire RAID;
  - Chargeurs automatiques de disques :
    - Densité optique des supports;
- Application d'imagerie :
  - Visualisateurs;
  - Gestion de base d'imagerie;

- Index, etc.

Comme on pouvait s'y attendre, il était impossible de faire respecter des niveaux précis de rendement en production par le sous-traitant, et ce, à cause de toutes les conditions ambiantes qui restaient indépendantes de sa volonté. Il demeurerait toutefois essentiel de veiller à ce que le rendement attendu du système du sous-traitant ne nuise pas à la réalisation du projet.

On a conçu des critères de rendement que l'on a intégrés à l'énoncé des critères d'acceptation de la DDP et mis en application. Dans un sous-réseau isolé, le sous-traitant a dû exécuter les tâches suivantes en employant la technologie de stockage et le serveur en production :

- mettre en lecture les versions du questionnaire du recensement de 1996;
- créer les images en reproduction multiple;
- garnir la base d'imagerie jusqu'à la moitié environ de la taille qu'elle aurait en production;
- à l'aide de deux postes de travail, extraire les quantités d'images représentatives prévues en production dans des délais précis; l'opération a été échelonnée (1, puis 10, puis 100 questionnaires, et ainsi de suite).

Dans ce scénario, le système du sous-traitant devait produire la première page d'un questionnaire en moins de 5 secondes et les pages suivantes en moins de 1 seconde.

Tout aussi importante était l'intégration de ce système aux systèmes de production créés à l'interne. Pour répondre aux critères, le sous-traitant a dû travailler avec un spécialiste de l'élaboration de systèmes et démontrer clairement que, avec l'application du recensement, on pouvait bien faire fonctionner un visualisateur (et produire des images).

À l'issue de ces essais, les représentants du secteur du recensement étaient raisonnablement sûrs que, si des problèmes de rendement devaient se poser, ce ne serait pas à cause du système du fournisseur.

Du côté «recensement» de la technologie, on a pris des mesures d'optimisation du rendement.

1. On a configuré, dimensionné et bien vérifié le réseau;
2. On a fait l'acquisition d'une grande antémémoire RAID pour y stocker les images fréquemment demandées;
3. On a acquis des Pentium avec assez de mémoire vive et d'espace disque;

4. On a aussi fait l'acquisition de cartes vidéo et de moniteurs à grand rendement.

Le résultat net a été une transmission satisfaisante d'images en production. Comme les temps d'arrêt ont été infimes, on a également constaté un haut degré de disponibilité du système.

**Facteurs de succès :** À l'époque où on a entrepris de se doter d'un système d'imagerie documentaire aux fins du recensement de l'agriculture, les exemples de catastrophes abondaient pour de tels projets, Statistique Canada ne comptait pas d'applications importantes d'imagerie et les gens chargés de mettre en place un tel système ne connaissaient absolument rien à cette technologie. L'exercice aurait pu se solder par un échec retentissant et fort coûteux pour le recensement canadien de l'agriculture.

On peut dégager un certain nombre de facteurs de réussite :

- **Organisation de Statistique Canada en informatique.** On avait sous la main pour le projet des compétences techniques en ordinateurs de gamme moyenne, en réseaux, en postes de travail et en élaboration de systèmes. Cette contribution a été aussi primordiale à l'étape de la production que dans les activités préparatoires.
- **Infrastructure technique de Statistique Canada :** Depuis quelque temps, l'organisme insistait fortement sur le respect des normes, ce qui a permis d'implanter dans cet environnement technique une technologie respectueuse des normes établies.
- **Spécialiste des méthodes :** Un spécialiste des méthodes a aidé les membres de l'équipe à analyser la technologie et à produire des statistiques utiles.
- **Liaison technique :** Aujourd'hui, les projets d'information font souvent appel à plusieurs disciplines techniques (ordinateurs de gamme moyenne, réseaux, postes de travail, élaboration d'applications, bases de données, produits, etc.). Dans ce projet d'imagerie, l'agent de liaison technique a coordonné les activités et représenté le secteur du recensement auprès de l'industrie et des groupes techniques.
- **Système clés en main :** Plutôt que de diviser un contrat entre plusieurs entreprises respectivement chargées de produire les images, de gérer le système d'imagerie et de fournir des services de visualisation, on a fait l'acquisition en bloc d'un système d'imagerie, d'où des garanties

d'interfonctionnalité de tous les éléments du système (ainsi, les images produites par les lecteurs seraient compatibles avec les visualisateurs). On prévenait en outre les problèmes de communication dans les interventions d'entretien. Pour le secteur du recensement, il n'y avait qu'un sous-traitant à qui s'adresser.

- **Simplicité :** L'absence de postes d'indexation et de contrôle de qualité a rendu l'application moins complexe. Ajoutons que l'existence d'une interface minimale entre l'application de production du recensement et le système d'imagerie FileNet a réduit les éléments et donc les problèmes d'intégration.
- **Expérience directe :** Il aurait été dangereux de s'en remettre à la documentation spécialisée et aux démonstrations, plus particulièrement en ce qui concerne la fonctionnalité des lecteurs automatiques. De plus, une évaluation pratique des moniteurs a permis de s'assurer d'une qualité suffisante des appareils. Il aurait été facile de trop dépenser ou encore de ne pas dépenser assez, ce qui aurait été une catastrophe. Dans la mesure du possible, le personnel du recensement a emprunté des modèles et en a fait l'expérience directe avec les documents et l'environnement du recensement.
- **Primauté des besoins :** Souvent dans ce projet, il aurait été plus facile de s'en remettre aux solutions classiques (postes d'indexation et de contrôle de qualité, etc.). De la sorte, on aurait inutilement augmenté les coûts, la complexité et les risques.
- **Persévérance :** Le projet d'imagerie aurait pu tourner court à plusieurs reprises.
- **Soutien et participation directe de la direction :** On sait fort bien qu'une foule de projets échouent par manque d'engagement de la direction. Le projet d'imagerie n'a pas souffert des conséquences d'un tel manque de participation des gestionnaires, loin de là. Le directeur adjoint et plus particulièrement le chargé de projet ont manifesté ouvertement leur appui, livré les batailles politiques et apporté une contribution de taille à la création du système.
- **Relations avec le sous-traitant :** Les rapports ont été excellents. La disponibilité, l'accessibilité et la disposition à faire des compromis ont été autant de facteurs de réussite. Il convient de noter que, avec ce projet, le sous-traitant s'est ouvert des portes dans l'administration fédérale. Le succès du projet était aussi important pour lui que pour le secteur du recensement, ce qui a sans doute joué dans les relations des deux parties.

- **Abondance des contributions** : Bien des gens ont été associés à la création du système d'imagerie. Les participants ont eu pour tâches :
  - d'évaluer les moniteurs, les cartes vidéo, les visualisateurs et les lecteurs;
  - de dimensionner le réseau et d'élaborer des programmes d'essai;
  - de définir la structure d'ensemble;
  - d'élaborer des DDP pour les essais en production et en projet pilote;
  - de définir comment les images seraient exploitées en production.
  - D'autres ont livré les inévitables batailles politiques.

Ce haut degré de participation a suscité un plus grand sentiment d'«appartenance» et explique sans doute en partie la réussite du projet.

- **Budget** : Les frais d'acquisition, d'implantation et d'entretien ne sont pas négligeables. Une solide caisse noire nous a permis de donner les coups de barre qui s'imposaient en période d'implantation.
- **Les questionnaires sur papier n'étaient plus une possibilité à envisager** : Souvent, lorsqu'on introduit des technologies, les systèmes techniques à remplacer sont toujours là. Même si on a d'impérieuses raisons de vouloir changer les choses, les gens continuent dans bien des cas à faire leur travail avec les outils qui leur sont familiers jusqu'à ce qu'ils aient l'occasion de voir la nouvelle technologie. Pour le recensement de l'agriculture, l'application de 1996 représentait un changement radical. Le système manuel de remise de questionnaires qui avait été employé en 1991 n'existait plus. En 1996, il a fallu des circonstances exceptionnelles pour que l'on se résigne à produire un document sur papier.
- **Acception** : Le personnel de production s'est montré largement satisfait de la qualité des images et de la fonctionnalité des visualisateurs. Il a jugé extrêmement productives les fonctions d'annotation, de défilement et de zoom.
- **Chance** : Le projet a profité de nombreux talents dans l'équipe et les services. Deux ou trois des grands artisans du projet étaient remarquables.

### *Section III : Réflexions et perspectives d'avenir*

On est parvenu au but fixé de transmission en temps utile d'images aux ordinateurs de bureau. Tous les utilisateurs et les observateurs ont bien accueilli -

c'est le moins que l'on puisse dire - la technologie. L'imagerie et un excellent jeu d'instruments de vérification et de validation se sont révélés une combinaison gagnante pour les utilisateurs. Tout ce que se demandait le personnel, c'est comment on avait pu faire par le passé pour traiter les données du recensement sans imagerie.

On a obtenu les avantages prévus et plus encore. L'imagerie s'est implantée dans une foule d'activités et d'applications où elle n'était pas du tout attendue. Voici quelques retombées imprévues :

- Le gain de productivité a largement dépassé les 10 % prévus. Dans un grand nombre d'opérations par commis, le visualisateur a automatiquement fait un zoom sur la partie problème du questionnaire et la correction s'est faite en direct. On travaillait bien plus vite qu'en feuilletant des questionnaires et des imprimés d'ordinateur. Dans le système manuel, il fallait notamment écrire la correction sur l'imprimé pour que quelqu'un d'autre l'apporte ensuite en direct.
- En période de collecte, il est possible de recevoir plusieurs questionnaires de recensement d'une exploitation agricole. Il y a une marche à suivre pour constater cette réception multiple et prendre des mesures correctives. En consultant des images des questionnaires d'une même exploitation sur écran partagé, on épargne bien du temps.
- Dans cette même période de collecte, les problèmes de sous-dénombrement obligeaient les préposés à identifier les exploitations oubliées ou manquantes. Pour ce faire, il fallait comparer les listes du recensement de 1996, des services fiscaux, des enquêtes et des recensements antérieurs. Grâce aux images, les vérificateurs ont pu rapidement voir si une exploitation figurait dans la base de données. On a réduit et les coûts de suivi sur le terrain et le fardeau de réponse et les risques de dédoublement des questionnaires. Les images ont aussi donné des indications (descriptions des terres, etc.) qui ne comptaient pas parmi les données saisies dans la base de données structurées.
- Un service en croissance rapide dans le secteur canadien de l'agriculture est celui des élevages porcins à contrat. Nombre d'éleveurs à contrat n'ont pas déclaré leurs porcins sur leur questionnaire malgré les instructions données. Comme les porcs ne leur appartenaient pas, ils n'ont pas jugé bon de les inclure. Nous avons

donc des données sur les exploitations, mais sans les porcs. En comparant le recensement de 1996 à l'enquête semestrielle sur les porcins, nous avons pu voir quels exploitants indiquaient un grand nombre de porcs dans l'enquête, mais non pas dans le recensement. Les animaux manquaient-ils par erreur de traitement ou de réponse? On a découvert la solution en consultant rapidement les images. On a ensuite fait le suivi nécessaire auprès des recensés pour qu'ils ajoutent les porcs manquants.

- Avec les méthodes d'imputation automatisées, il semble toujours se présenter des situations où, pour chaque module de questions, il est impossible de trouver un enregistrement donneur. Fréquemment, des erreurs de réponse ou de traitement sont en cause. Grâce aux images, les agents spécialisés ont pu prendre des mesures correctives. Dans les recensements du passé, on apportait les corrections au jugé pour faire passer les questionnaires à la prochaine étape d'imputation. Si on avait pris le temps de repérer et d'extraire les documents sur papier, on aurait sérieusement retardé les opérations d'imputation.

L'imagerie a-t-elle fait ses frais en une seule période de recensement?

Non, et il ne fallait pas s'y attendre. Comment mettre une valeur économique sur la qualité des données, la satisfaction au travail ou le coût du report d'une date de diffusion?

Pourra-t-on amortir les frais du système d'imagerie à court terme (en cinq ans)? Oui, sûrement. Le matériel continue à être utilisé au recensement dans des essais, ainsi que dans bien d'autres projets en expérimentation ou en production. Il jouera aussi un rôle - avec des perfectionnements et des compléments - dans le recensement de 2001.

L'expérience fructueuse qui en a été faite au recensement de l'agriculture a encouragé d'autres secteurs à essayer et à implanter non seulement la technologie d'imagerie, mais aussi d'autres applications RIC. Voici des secteurs qui s'occupent aujourd'hui de cette technologie :

- Projet des données complètes sur les exploitations agricoles;
- Recensement de la population;
- Division des données fiscales;
- Division de l'organisation et des finances de l'industrie;
- Division des opérations et de l'intégration;

- Division des services de soutien administratif (Gestion des documents);
- Projet d'amélioration des statistiques économiques provinciales (PASEP).

Parfois, le recensement de l'agriculture s'est chargé de travaux se rattachant à d'autres projets. Ce travail supplémentaire :

- aide à payer le système et permet de partager les frais d'entretien et de licence;
- soutient des recherches liées;
- entretient et enrichit l'expérience du personnel;
- ajoute des caractéristiques et des capacités au système.

Nul doute que la décision d'implanter un système d'imagerie a été la bonne. Le risque que l'on a pris était acceptable, mais non négligeable. S'il y avait eu de nouvelles difficultés, tout le projet aurait été en danger.

L'expérience prouve qu'il valait la peine de prendre un tel risque. Somme toute, cet élément de risque a joué un rôle primordial en aidant le secteur du recensement à respecter ses dates prévues de diffusion.

Divers événements avaient retardé les activités. À eux seuls, les problèmes de sous-dénombrement avaient fait accuser un retard de plusieurs semaines aux tâches de validation des données. De plus, les problèmes du système généralisé d'imputation avaient imposé une grande opération de colmatage et une réimputation de toutes les variables financières pour deux des principales provinces du pays.

Par d'exceptionnels efforts concertés des membres de l'équipe et grâce aux gains d'efficacité permis par le système d'imagerie, les données du recensement ont pu être diffusées à temps.

Où la technologie mène-t-elle le recensement de l'agriculture de 2001? La technologie RIC peut-elle traiter des questionnaires de 16 pages sur papier ministre remplis par 300 000 exploitants agricoles qui gagnent en âge?

Là encore, on observe une diversité d'opinions et de camps. Il y a des membres de l'équipe de projet qui étaient conservateurs en 1996 et qui sont progressistes aujourd'hui et vice versa. D'autres choses ont changé aussi. Nous savons que l'imagerie fonctionne et donne de bons résultats. Nous avons plus d'expérience et de confiance, et plus de temps aussi.

La recherche RIC pour le recensement de l'agriculture a débuté il y a presque un an, soit quatre ans avant le recensement de 2001. Les premiers résultats sont certes encourageants, mais il est bien trop tôt pour établir des prévisions et encore plus pour prendre des décisions. C'est en 1998 qu'on commencera à évaluer la réaction des recensés aux questionnaires adaptés à la technologie RIC.

Quelle sera notre position de repli si la RIC est jugée peu rentable ou peu appropriée au recensement de l'agriculture? Une possibilité «conservatrice» serait l'imagerie avec plus de traitement en parallèle. Un autre grand pas en avant serait la «lecture automatique de marques») avec extraction-images.

Certains voient dans l'imagerie documentaire une technologie transitoire dans la voie menant au bureau sans papier. On peut se demander quand la déclaration électronique remplacera la déclaration sur papier aux fins du recensement de l'agriculture.

### **BIBLIOGRAPHIE**

1. Bradshaw, C., et Duggan, J. (1997) Évaluation de l'imagerie documentaire appliquée au recensement canadien de l'agriculture de 1996.
2. Duggan, J. (1996) Electronic Imaging in Support of the 1996 Census of Agriculture, document présenté à la conférence internationale sur la collecte assistée par ordinateur des données des enquêtes.
3. Green, I. (1995) Questionnaire Imaging Pilot, rapport interne, Statistique Canada.



## *Serie de documents de travail sur l'agriculture et le milieu rural*

La Division de l'agriculture publie des documents de travail sur la recherche, les résultats analytiques, les techniques statistiques, les méthodes et les concepts. Pour commander, appelez la Division de l'agriculture sans frais en composant le 1-800-465-1991. (Certains des documents de travail sont disponibles sur Internet.)

N° 1	(21-601-MPF80001)	<b>Description de la méthode Theil de prévision de l'erreur quadratique moyenne pour la statistique agricole (1980)</b>	Stuart Pursey
N° 3	(21-601-MPF81003)	<b>Examen du Projet de l'estimation du bétail et recommandations de mesures à prendre (1981)</b>	Bernard Rosien et Elizabeth Leckie
N° 4	(21-601-MPF84004)	<b>Le secteur canadien des oléagineux : vue d'ensemble (1984)</b>	Glenn Lennox
N° 5	(21-601-MPF84005)	<b>Analyse préliminaire de la contribution des paiements directs du gouvernement dans le revenu agricole net réalisé (1984)</b>	Lambert Gauthier
N° 6	(21-601-MPF84006)	<b>Les caractéristiques des exploitants entrant en agriculture et leurs entreprises au sud de l'Ontario pour la période 1966 à 1976 (1984)</b>	Jean B. Down
N° 7	(21-601-MPF84007)	<b>Sommaire des programmes d'aide à la production agricole aux États-Unis (1984)</b>	Allister Hickson
N° 8	(21-601-MPF84008)	<b>Intensité de la pratique de la jachère dans les Prairies : Une analyse des données du recensement de 1981 (1984)</b>	Les Macartney
N° 9	(21-601-MPF85009)	<b>Évolution de la structure du secteur porcin au Canada (1985)</b>	Mike Shumsky
N° 10	(21-601-MPF86010)	<b>Révisions au traitement des loyers de maisons imputés dans les comptes de fermes canadiennes, 1926-1979 (1986)</b>	Mike Trant
N° 11	(21-601-MPF92011)	<b>L'estimateur par le quotient : explication intuitive et utilisation pour estimer les variables agricoles (1992)</b>	François Maranda et Stuart Pursey
N° 12	(21-601-MPF91012)	<b>L'effet de la distorsion géographique causée par la règle de l'emplacement (1991)</b>	Rick Burroughs
N° 13	(21-601-MPF91013)	<b>La qualité des données agricoles : forces et faiblesses (1991)</b>	Stuart Pursey
N° 14	(21-601-MPF92014)	<b>Autres cadres d'examen des données rurales (1992)</b>	A.M. Fuller, Derek Cook et Dr John Fitzsimons
N° 15	(21-601-MPF93015)	<b>Tendances et caractéristiques relatives aux régions rurales et aux petites villes du Canada (1993)</b>	Brian Biggs, Ray Bollman et Michael McNames
N° 16	(21-601-MPF92016)	<b>La microdynamique et l'organisation économique de la famille agricole dans le changement structurel en agriculture (1992)</b>	Phil Ehrensaft et Ray Bollman
N° 17	(21-601-MPF94017)	<b>Consommation de céréales et de graines oléagineuses par le bétail et la volaille, Canada et provinces, 1992</b>	Section du bétail et des produits d'origine animale
N° 18	(21-601-MPF94018)	<b>Changements structurels dans le domaine agricole - Étude comparative des tendances et des modèles observés au Canada et aux États-Unis</b>	Ray Bollman, Leslie A. Whitener et Fu Lai Tung



N° 19	(21-601-MPF94019)	<b>Revenu total de la famille agricole selon le type d'exploitation et la taille de celle-ci, et selon la région, en 1990 (1994)</b>	Saiyed Rizvi, David Culver, Lina Di Piéto et Kim O'Connor
N° 20	(21-601-MPF94020)	<b>L'adaptation dans le secteur agricole au Canada (1994)</b>	George McLaughlin
N° 21	(21-601-MPF93021)	<b>Microdynamique de la croissance et de la décroissance des exploitations agricoles : une comparaison Canada - États-Unis</b>	Fred Gale et Stuart Pursey
N° 22	(21-601-MPF92022)	<b>Les structures des gains des ménages agricoles en Amérique du Nord – Positionnement pour la libéralisation des échanges</b>	Leonard Apedaile, Charles Barnard, Ray Bollman et Blaine Calkins
N° 23	(21-601-MPF92023)	<b>Secteur de la pomme de terre : comparaison entre le Canada et les États-Unis</b>	Glenn Zepp, Charles Plummer et Barbara McLaughlin
N° 24	(21-601-MPF94024)	<b>Étude comparative des données américaines et canadiennes sur la structure des fermes</b>	Victor J. Oliveira, Leslie A. Whitener et Ray Bollman
N° 25	(21-601-MPF94025)	<b>Méthodes statistiques de la Sous-section de la commercialisation des céréales, document de travail, version 2</b>	Karen Gray
N° 26	(21-601-MPF94026)	<b>Rendement des exploitations agricoles : Estimations établies à partir de la base de données complètes sur les exploitations agricoles</b>	W. Steven Danford
N° 27	(21-601-MPF94027)	<b>La mesure de l'emploi touristique dans les régions rurales</b>	Brian Biggs
N° 28	(21-601-MPF95028)	<b>Délimitation de l'écoumène agricole canadien de 1991</b>	Timothy J. Wershler
N° 29	(21-601-MPF95029)	<b>Étude cartographique de la diversité des économies rurales : une typologie préliminaire du Canada rural</b>	Liz Hawkins
N° 30	(21-601-MPF96030)	<b>Structure et tendances de l'emploi rural au Canada et dans les pays de l'OCDE</b>	Ron Cunningham et Ray D. Bollman
N° 31	(21-601-MPF96031)	<b>Une nouvelle approche pour les régions autres que les RMR/AR</b>	Linda Howatson-Leo et Louise Earl
N° 32	(21-601-MPF96032)	<b>L'emploi dans l'agriculture et ses industries connexes en région rurale : structure et changement 1981-1991</b>	Sylvain Cloutier
N° 33	(21-601-MPF98033)	<b>Exploiter une ferme d'agrément – pour le plaisir ou le profit?</b>	Stephen Boyd
N° 34	(21-601-MPF98034)	<b>Utilisation de la technologie d'imagerie documentaire dans le recensement canadien de l'agriculture de 1996</b>	Mel Jones et Ivan Green
N° 35	(21-601-MPF98035)	<b>Tendances de l'emploi au sein de la population active non métropolitaine</b>	Robert Mendelson
N° 36	(21-601-MPF98036)	<b>La population des milieux ruraux et des petites villes s'accroît pendant les années 90</b>	Robert Mendelson et Ray D. Bollman
N° 37	(21-601-MPF98037)	<b>La composition des établissements commerciaux dans les petites et les grandes collectivités du Canada</b>	Robert Mendelson
N° 38	(21-601-MPF98038)	<b>Le travail hors ferme des exploitants de fermes de recensement : Aperçu de la structure et profils de mobilité</b>	Michael Swidinsky, Wayne Howard et Alfons Weersink

N° 39	(21-601-MPE99039)	<b>Le capital humain et le développement rural : quels sont les liens?</b>	Ray D. Bollman
N° 40	(21-601-MPE99040)	<b>Utilisation de l'ordinateur et d'Internet par les membres des ménages ruraux</b>	Margaret Thompson-James
N° 41	(21-601-MPE99041)	<b>Les Cotisations aux réer des producteurs agricoles Canadiens en 1994</b>	Marco Morin
N° 42	(21-601-MPE99042)	<b>Intégration des données administratives et des données d'enquête de recensement</b>	Michael Trant et Patricia Whitridge

<b><u>Prix:</u></b>	<b>1 copie</b>	<b>Gratuite</b>
	<b>2 à 10 copies</b>	<b>50,00 \$</b>
	<b>11 à 25 copies</b>	<b>75,00 \$</b>
	<b>26 à 38 copies</b>	<b>100,00\$</b>





# BON DE COMMANDE

## Statistique Canada

### POUR COMMANDER :

#### COURRIER

Division de l'agriculture  
Statistique Canada  
12<sup>ème</sup> étage,  
Édifice Jean Talon  
Parc Tunney  
Ottawa (Ontario)  
K1A 0T6

#### TÉLÉPHONE

**1 800 465-1991**  
Faites débiter votre compte  
VISA ou MasterCard. De  
l'extérieur du Canada et  
des États-Unis et dans la  
région d'Ottawa, composez  
le (613) 951-7134. Veuillez  
ne pas envoyer de  
confirmation.

#### TÉLÉCOPIEUR

**1 613 951-1680**  
VISA, MasterCard et bon de  
commande seulement.  
Veuillez ne pas envoyer  
de confirmation. Le bon  
télécopié tient lieu de  
commande originale.

#### INTERNET [agriculture@statcan.ca](mailto:agriculture@statcan.ca)

(Veuillez écrire en majuscules)

#### 1 800 363-7629

Appareils de  
télécommunications  
pour les malentendants

Compagnie _____		Service _____	
À l'attention de _____		Fonction _____	
Adresse _____			
Ville _____		Province _____	
Code Postal _____		Téléphone _____	Télécopieur _____
Adresse de courrier électronique: _____			

### MODALITÉS DE PAIEMENT :

(Cochez une seule case)

Veuillez débiter mon compte     VISA     MasterCard

\_\_\_\_\_ N° de carte

\_\_\_\_\_ Date d'expiration

\_\_\_\_\_ Détenteur de carte (en majuscules s.v.p.)

\_\_\_\_\_ Signature

Paiement inclus \$ \_\_\_\_\_

N° du bon de commande \_\_\_\_\_  
(veuillez joindre le bon)

\_\_\_\_\_ Signature de la personne autorisée

N° du document de travail	Titre	Numéro au catalogue	N° de copies	Total \$
<b>► PRIX:</b> 1 copie                    Gratuite 2 à 10 copies            50,00 \$ 11 à 25 copies           75,00 \$ 26 à 38 copies           100,00 \$		<b>TOTAL (encrer Cnd. ou U.S. s.v.p.)</b> \$ Cnd./ U.S.    \$		
<b>► Les prix peuvent être modifiés sans préavis. Pour vérifier les prix courants, veuillez composer le 1 800 465-1991.</b>		<b>Réduction (s'il y a lieu)</b>		
<b>► Veuillez noter que les prix au catalogue pour les clients de l'extérieur du Canada sont donnés en dollars américains. Les clients de l'extérieur du Canada paient le montant total en dollars US tirés sur une banque américaine.</b>		<b>TPS (7 %) (Clients canadiens seulement, s'il y a lieu)</b>		
<b>► Les clients canadiens paient en dollars canadiens et ajoutent soit la TPS de 7 % et la TVP en vigueur, soit la TVH.</b>		<b>TVP en vigueur (Clients canadiens seulement, s'il y a lieu)</b>		
<b>► Le chèque ou mandat-poste doit être établi à l'ordre du Receveur général du Canada.</b>		<b>TVH en vigueur (N.-É., N.-B., T.-N.)</b>		
<b>► TPS N° R121491807</b>		<b>TOTAL GÉNÉRAL</b>		\$
		<b>PF097177</b>		



## MERCI POUR VOTRE COMMANDE!