

Distr.
GÉNÉRALE

CES/SEM.47/23
16 janvier 2002

FRANCAIS
ANGLAIS et FRANCAIS seulement

COMMISSION STATISTIQUE et
COMMISSION ÉCONOMIQUE POUR L'EUROPE

COMMISSION DES COMMUNAUTÉS
EUROPÉENES

CONFÉRENCE DES STATISTICIENS
EUROPÉENS

EUROSTAT

Séminaire commun CEE-Eurostat sur les
Systèmes intégrés d'information statistique
et les questions connexes (ISIS 2002)
(Genève, Suisse, 17-19 avril 2002)

Point IV: Moyens permettant aux systèmes d'information statistique de mieux répondre aux besoins des utilisateurs

LES MÉTADONNÉES À STATISTIQUE CANADA

Communication sollicitée

Rapport envoyé par Statistique Canada ¹

I. INTRODUCTION

1. Depuis quelques années, on assiste à une prolifération de centres et d'entrepôts de données, et le mûrissement de la technologie facilite maintenant l'accessibilité aux données par Internet. Les utilisateurs de données ont l'énorme tâche de filtrer l'information, d'évaluer la comparabilité des sources de données disparates puis de déterminer la pertinence et les qualités des données dont ils ont besoin. Un producteur de données fait face à des tâches du même ordre : « Existe-t-il des données qui peuvent fournir une valeur supplémentaire ou qui simplifient le processus de production? » « Quelqu'un a-t-il déjà produit des données identiques ou similaires qui me permettraient d'en comparer la qualité? » Même s'il y a des moteurs de recherche et des outils d'analyse que l'utilisateur peut utiliser pour faciliter les recherches, l'analyse et la présentation des données, ce sont les métarenséignements et les métadonnées² qui confèrent signification et contexte aux données.

¹ Préparé par Amie Lee.

² Les métarenséignements s'entendent des concepts et définitions qui décrivent les procédés d'affaires servant à la création des données. Les métarenséignements sont utilisés pour évaluer et comprendre les données d'enquête. Les métadonnées désignent la description des données physiques. Les métadonnées favorisent l'accès aux données d'enquête et appuient l'échange de données entre les systèmes et les organisations. On utilisera le terme *métadonnées* pour désigner à la fois les métarenséignements et les métadonnées.

2. Le présent document aborde l'évolution des métarensignements et des métadonnées à Statistique Canada, les outils logiciels utilisés pour l'élaboration de la Base de métadonnées intégrées (BMDI), les rôles actifs et passifs de l'information, l'intégration de la BMDI avec les systèmes existants et leurs procédés administratifs, et enfin les voies d'accès à l'information stockée dans la BMDI.

3. La BMDI est un registre central de métarensignements et de métadonnées qui sont nécessaires pour décrire, répertorier, analyser et classer clairement des renseignements et des données statistiques à Statistique Canada.

II. HISTORIQUE DES MÉTARENSEIGNEMENTS ET DES MÉTADONNÉES

4. Il existe pour les données de Statistique Canada des *métadonnées* complètes et de bonne qualité, cependant elles sont réparties sur divers supports distincts, par exemple, les spécifications de conception d'un outil de génie logiciel assisté par ordinateur, des systèmes d'échantillonnage, de contrôle ou d'imputation, des bases de données et des entrepôts de données, des répertoires de données ou sous forme d'une publication ou d'un document.

5. Il y a eu à Statistique Canada diverses initiatives destinées à créer un entrepôt de *métadonnées*. Le Système de documentation des données statistiques (SDDS) a été la première initiative du genre. Ce système renfermait des renseignements descriptifs au sujet des enquêtes réalisées à Statistique Canada. La Division des normes a mis à jour chaque année les renseignements des enquêtes en communiquant avec les gestionnaires de programme de chacun des secteurs de production de données. Le contenu du Système de documentation des données statistiques n'a pas réussi à fournir des métarensignements opportuns et n'a pas répondu aux besoins de l'organisme en matière de métadonnées.

6. Dans les années 90, on a élaboré un outil thématique de recherche, fonctionnant avec Folio Views, dont on a assuré la mise à jour centralisée pour les statistiques sociales. Ce système a fourni un bon accès thématique aux métarensignements pour les statistiques sociales, cependant il ne présentait pas l'utilisation courante du vocabulaire ou des expressions et la mise à jour de la classification thématique nécessitait un traitement manuel et une maintenance énormes pour la collecte et l'extraction des données. Toujours dans les années 90, on a mis au point le modèle de collecte électronique Paradox pour atténuer une partie des problèmes de collecte et de mise à jour de l'outil thématique de recherche. Cet outil a été déployé uniquement pour les secteurs spécialisés de la statistique sociale, et sa mise à jour annuelle incombait aux gestionnaires de programme. Le modèle de collecte électronique Paradox n'autorisait pas l'échange de données parce que chaque secteur spécialisé possédait sa propre version de la base de données. La plateforme matérielle, les outils logiciels et la conception de la base de données utilisés pour confectionner le modèle de collecte n'étaient pas suffisamment robustes pour justifier leur utilisation partout dans l'organisme.

7. En outre, dans les années 90, la Division des services d'accès et de contrôle des données tenait à jour le Meta inventaire des acquis statistiques qu'elle avait élaboré. Ce système servait à entreposer les renseignements d'archives pour tous les fichiers maîtres de données statistiques épurées de Statistique Canada. La Division des services d'accès et de contrôle des données mettait l'information à jour tous les ans en communiquant avec les gestionnaires de programme des secteurs spécialisés. Les données ont satisfait aux exigences des archivistes des Archives nationales.

8. Chacun des systèmes décrits ci-avant recueillait séparément ses propres *métadonnées*. Outre son incohérence d'un système à l'autre, l'information imposait un fardeau de réponse aux gestionnaires de programme, car chaque système recueillait les mêmes renseignements mais à différents moments pendant l'année.

9. En 1998, on créait la base de métadonnées intégrées afin d'établir un entrepôt de *métadonnées* pour tout Statistique Canada. Le projet visait à définir les besoins en matière de *métadonnées* et à regrouper les diverses sources de métareenseignements et de métadonnées dans un répertoire central pour toutes les activités statistiques de l'organisme. Le but était de réduire au minimum la duplication de l'information entre les divers systèmes, d'offrir un meilleur accès à l'information tant aux clients internes qu'externes, d'éliminer le fardeau de réponse des secteurs spécialisés et d'aider à l'harmonisation des concepts et des variables. L'entrepôt et le système de la base de métadonnées intégrées (BMDI) sont actuellement en production et les mises à jour sont limitées à celles de la Division des normes³.

III. ÉLABORATION DU CONTENU

10. Pour élaborer le contenu de la BMDI (métareenseignements), nous nous sommes inspirés des documents suivants :

- politique de Statistique Canada visant à informer les utilisateurs sur la qualité des données et la méthodologie;
- lignes directrices de Statistique Canada concernant la qualité;
- la politique concernant les normes de Statistique Canada;
- Statistique Canada, Un aperçu du traitement des enquêtes auprès des entreprises, avril 1999;
- le modèle d'entrepôt des métadonnées statistiques du Bureau of Census des États-Unis (MDR)⁴.

11. Entre juin 1998 et septembre 1999, des consultations ont eu lieu avec la Division de la diffusion, la Division des services d'accès et de contrôle des données, des représentants du Projet d'amélioration des statistiques économiques provinciales, des représentants de l'Initiative de démocratisation des données, la Division des méthodes d'enquête auprès des entreprises, la Division des méthodes d'enquêtes sociales, la Division des normes, la Bibliothèque et centre d'information, divers secteurs spécialisés et la Division du développement de systèmes⁵. Pendant ces séances de consultation, les participants ont été confrontés au problème de la date de mise en œuvre. Comme les secteurs spécialisés avaient établi leurs propres exigences en matière de *métadonnées* pour leurs calendriers de production, ils ont mis au point leurs propres systèmes de *métadonnées* pendant l'étape de conception et d'élaboration de la BMDI. Les *métadonnées* de ces systèmes seront abordées à une date ultérieure⁶.

12. En juin 2001, nous nous sommes inspirés des composantes suivantes pour élaborer le contenu en métadonnées de la BMDI :

- entrepôt de métadonnées statistiques du Bureau of Census des États-Unis (MDR);
- ISO/IEC CD 11179-3, Information Technology – Data Management and Interchange – Metadata Registries (MDR) – Part 3: Registry Metamodel (MDR3), June 6, 2001-12-11;
- ISO/IEC PDTR 20943-3, Information Technology – Achieving metadata registry (MDR) content consistency – Part 3: Value domains, February 26, 2001-12-11;
- Neuchâtel Group, The Neuchâtel Terminology: Classification Database Object Types and their Attributes.
- Statistique Canada, Un aperçu du traitement des enquêtes auprès des entreprises, avril 1999.

³ On prévoit que le système sera déployé à la grandeur de l'organisme dès qu'on aura terminé les modules d'accès et de sécurité des utilisateurs, d'inscription et de mise à jour des versions.

⁴ Dan Gillman a envoyé le modèle par courrier électronique.

⁵ Une communication préparée par Paul Johannis présente un aperçu du contenu de la BMDI. Le document est intitulé « Statistics Canada's Integrated Meta Database Current Status and Future Plans », Séance de travail sur METIS, 28-30 novembre 2000, Washington (D. C.).

⁶ Les *métadonnées* de ces systèmes pourraient représenter une source de données pour le chargement de la BMDI, ou la BMDI remplacerait entièrement ces systèmes. On étudiera la question dès qu'on aura défini tous les procédés administratifs d'interaction avec la BMDI.

13. La Division des normes, la Division du développement de systèmes et Culture, tourisme et Centre de la statistique de l'éducation⁷ ont examiné les exigences en métadonnées et étendu le modèle de la BMDI pour assurer une concordance entre les domaines de valeur et pour fournir des fonctions automatisées de codage à venir pour le Système de codage pour la classification des industries de Statistique Canada (SCCI).

IV. PROCESSUS DE COLLECTE ET OUTILS DE COLLECTE

14. Des écrans de collecte de la BMDI et des utilitaires de chargement spéciaux ont été élaborés pour l'activité de collecte et de planification⁸. La collecte et la vérification des *métadonnées* représentent une tâche imposante à cause de la quantité même des données nécessaires pour décrire l'activité d'enquête et ses produits de données. Même si la collecte et la mise à jour représentent actuellement un processus central confié à la Division des normes, on a cerné, dans le processus de diffusion existant, des déclencheurs de mise à jour des *métadonnées*. L'intégration de la collecte des *métadonnées* dans le flux de travail existant représente une collecte passive de *métadonnées*, cependant elle accroît la probabilité de disposer de *métadonnées* à jour puisque ces déclencheurs agissent comme points de contrôle des *métadonnées*. La collecte active de *métadonnées* deviendra une réalité lorsque les *métadonnées* joueront un rôle plus actif dans la planification, l'élaboration et la gestion des activités d'enquête et des produits d'enquête.

A. Chargement des métareenseignements

15. On a intégré dans la BMDI les métareenseignements provenant des systèmes suivants, qu'on a ensuite mis au rancart :

- Système de documentation des données statistiques;
- Outil thématique de recherche;
- Modèle de collecte électronique Paradox;
- Meta inventaire des acquis statistiques;
- Système de répertoire des questionnaires⁹.

On a regroupé et chargé dans la BMDI les métareenseignements de chacun de ces systèmes en utilisant des outils de conversion spécialement élaborés à cette fin en Access 97 pour la saisie et l'analyse, et des procédures en PL/SQL d'Oracle pour le chargement des données dans la BMDI. La Division des normes gère l'activité de conversion, le regroupement de l'information et la vérification des données chargées. On a mis au rancart les outils en Access 97 et les procédures en PL/SQL d'Oracle une fois terminé le chargement dans la BMDI. On a aussi recueilli des métareenseignements des publications, qu'on a ensuite saisies à la main par l'entremise des écrans de collecte de la BMDI.

16. La vérification a exigé beaucoup de temps et de ressources, pour les raisons suivantes :
- Le chevauchement du contenu provenant des systèmes mis au rancart nécessitait une vérification [et une approbation de chacun des gestionnaires de système et gestionnaires de programme des secteurs spécialisés.]
 - L'information doit être disponible dans les deux langues officielles¹⁰ et les descriptions textuelles doivent correspondre dans les deux langues officielles.

⁷ Les représentants provenaient du programme d'éducation qui participe, en tant que programme pilote, à l'élaboration du contenu de l'éducation.

⁸ En décembre 2001, près de 80 % des écrans de saisie avaient été livrés à la Division des normes pour fins de production.

⁹ Le Système de répertoire des questionnaires se compose des questionnaires utilisés pendant chaque exercice. La Division des normes tient à jour l'inventaire.

¹⁰ Français et anglais

17. On a simplifié le processus de vérification en recourant aux procédés et véhicules existants de diffusion des métareenseignements. La Division des normes avait déjà publié le contenu du SDDS sur l'intranet de Statistique Canada ainsi que sur Internet au moyen de pages HTML produites à la main. Les gestionnaires de programme ont utilisé ces pages pour revoir les données et envoyer les mises à jour à la Division des normes pour intégration dans le système SDDS. On a ensuite remplacé ces pages HTML produites à la main par des pages HTML statiques produites automatiquement par la BMDI, tandis que le processus de mise à jour par les gestionnaires de programme restait inchangé.

B. Chargement des métadonnées

18. Le chargement initial des métadonnées dans la BMDI est limité aux variables de sortie diffusées sur CANSIM. Ce chargement exige aussi beaucoup de ressources, car il nécessite l'élaboration d'un contenu normalisé et harmonisé qui repose sur un examen de toutes les variables CANSIM et de leurs définitions. L'harmonisation s'effectue en définissant des significations types aux valeurs puis en déterminant l'ensemble de valeurs permises liées aux significations types des valeurs. L'élaboration du contenu comporte la définition de classes d'objets, propriétés, concepts d'éléments de données, domaines conceptuels, domaines de valeurs et éléments de données pour les variables de sortie du CANSIM. La Division des normes est chargée d'examiner et d'élaborer ce contenu. Le processus de vérification exige beaucoup de ressources humaines pour les mêmes raisons que nous avons décrites dans le cas du chargement des métareenseignements.

19. Le chargement, la mise à jour et la vérification centralisés par la Division des normes comporte plusieurs avantages :

- conformité aux politiques et directives de Statistique Canada;
- le contenu de l'information est uniforme pour toutes les *métadonnées* d'un secteur spécialisé à l'autre;
- la qualité de l'information est uniforme pour toutes les *métadonnées* d'un secteur spécialisé à l'autre;
- les fournisseurs de *métadonnées* disposent d'un exemple concret du niveau de détail souhaité;
- les fournisseurs de *métadonnées* disposent d'un exemple concret du style de produits à livrer;
- les fournisseurs de *métadonnées* disposent d'un point de contact, ce qui réduit le temps qu'ils consacrent au processus;
- l'harmonisation est réalisable.

20. En contrepartie, la Division des normes a l'énorme tâche de saisir et de gérer l'information, outre l'analyse et l'élaboration du contenu. Divers outils ont été conçus pour l'aider dans cette tâche.

21. On a élaboré un chargeur/exportateur à l'aide de Visual Basic 5.0 et de XML-DBMS version 1.01 afin de créer une solution basée sur XML qui pourrait charger et exporter les classifications types (comme le SCIAN)¹¹.

22. On a élaboré un outil de collecte en Access97 afin de réunir les métadonnées de CANSIM de manière à ce que la Division des normes puisse analyser et préparer les métadonnées avant de les charger

¹¹ XML-DBMS est un système servant à transférer des données entre des documents XML et des bases de données relationnelles. Le système affiche un document XML sous forme d'arbre d'objets puis utilise une fonction de mappage objet-relation pour mapper les objets à une base de données relationnelles. Cette version de l'importateur/exportateur sert au chargement initial des données de classification dans la BMDI. La décision de charger et d'exporter uniquement les classifications types reposait sur les classifications types utilisées dans CANSIM et le SCCI. En décembre 2001, le chargeur/exportateur de la BMDI était à l'étape de vérification. On s'attend que le chargement des classifications types dans la BMDI sera terminé en mars 2002. On prévoit examiner et étendre dans l'avenir la fonctionnalité d'importation/exportation de la BMDI à d'autres composantes.

dans la BMDI. On a retenu une solution en Access afin d'exploiter les connaissances acquises en Access. On mettra cet outil au rancart dès que les métadonnées de CANSIM seront chargées dans la BMDI.

C. Écrans de collecte de la BMDI

23. La conception des écrans de collecte de *métadonnées* visait à :

- répartir la quantité d'information en éléments maniables destinés à être livrés par modules aux fins de la vérification et de la production;
- faciliter la collecte en optimisant la quantité d'information sur un écran tout en fournissant suffisamment d'information pour une collecte intuitive de la part des secteurs spécialisés;
- réduire le temps d'élaboration;
- réduire le temps nécessaire à l'achèvement des essais de système;
- réduire l'incidence des changements sur le système à la suite des nouvelles exigences;
- réduire le temps de formation nécessaire à l'utilisation du système de saisie.

24. On a intégré la rétroaction reçue des utilisateurs du modèle de collecte électronique Paradox aux écrans de collecte de la BMDI. Voici les questions abordées :

- les utilisateurs ont le choix entre une interface française ou anglaise avec affichage simultané du français et de l'anglais¹²;
- fournir une solution de rechange à une collecte du type question-réponse¹³;
- accès efficace aux secteurs d'entrée sans imposer un défilement excessif;
- capacité d'adapter les étiquettes et les questions.

25. La méthodologie orientée-objet a permis de satisfaire les objectifs de conception du système de saisie. Toutes les composantes administrées sont gérées de la même façon. Du point de vue de la collecte, les écrans de recherche et les écrans de saisie pour toutes les composantes administrées présentent une interface et un comportement identiques, à l'exception du contenu de l'information qui variait selon le type de composantes administrées. Cela a permis de réduire au minimum la formation requise à l'utilisation du système de saisie. Dès qu'un utilisateur connaissait les fonctions d'entrée d'une composante administrée, seul le contenu de l'information changeait pour les autres composantes administrées. Le temps d'essai est réduit, car lorsqu'une composante administrée est vérifiée, seuls les changements de contenu doivent faire l'objet d'un essai dans les autres composantes administrées. Du point de vue du développement des systèmes, le temps de développement diminue pour chaque nouvelle composante administrée. La classe Administered Component gère l'aspect administratif¹⁴ et tous les objets du type de la composante administrée sont des sous-classes de la classe Administered Component de qui ils héritent les classes et les méthodes. Les concepteurs n'ont qu'à se concentrer sur les classes et méthodes propres à l'objet. Le temps de développement diminue pour chaque nouvelle composante administrée qui est ajoutée au système, cependant le temps initial de conception s'accroît, car le concepteur doit avoir une vue d'ensemble du système pour mettre au point une architecture qui utilisera efficacement les techniques orientées objet tout en exploitant la réutilisation des objets. Dans la mesure du possible, l'équipe de développement a appliqué des normes à toutes les étapes de la conception et de l'élaboration en utilisant des schémas de conception classiques orientés objet et des examens de conception et de code. On a utilisé le Unified Modelling Language (UML) pour établir le modèle de l'application de saisie. L'environnement de développement orienté objet est beaucoup plus technique et représente un changement complet de paradigme à Statistique Canada.

¹² Une interface française ou anglaise signifie que les directives, étiquettes de données, messages à l'utilisateur et messages d'aide apparaissent dans la langue choisie par l'utilisateur. Les zones de données texte en français et en anglais s'affichent toujours ensemble dans l'interface, et les fournisseurs de *métadonnées* doivent toujours fournir un texte français et anglais.

¹³ Dès que les utilisateurs connaissaient le système, la portion de l'écran utilisée par les questions est devenue encombrante et gênante pour la navigation.

¹⁴ Régions d'intendance, d'identification et de classification.

26. L'écran de collecte de la BMDI est une application client légère à deux paliers. Le palier de données et le palier de l'application se situent au même niveau grâce à la base de données relationnelles Oracle 8i et aux procédures stockées en PL/SQL. Le palier de présentation est une application JAVA¹⁵ élaborée à l'aide de Visual Age 3.0 d'IBM (JDK 1.1.7, Swing 1.0.3)¹⁶ à laquelle se greffent des pilotes JDBC d'Oracle pour les intergiciels. Les autres outils utilisés pour le développement comprennent System Architect 2001 V7.12, TOAD VI V6.2.10.29 (outil pour les concepteurs d'applications Oracle) pour le développement de PL/SQL et Jprobe Profiler and Memory Analyser.

V. DIFFUSION

27. La diffusion de la BMDI avait pour objectif d'appuyer la diffusion par l'entremise du site Web de Statistique Canada (www.statcan.ca). Le site Web offre un vaste éventail de renseignements et de statistiques qui s'adressent à une diversité d'utilisateurs de données.

A. Utilisation passive

28. Des pages HTML statiques sont produites automatiquement tous les jours par la BMDI pour l'intranet et Internet¹⁷. Les pages présentent des renseignements à jour sur le dernier cycle d'enquête¹⁸. Avec la mise en œuvre de la BMDI, le contenu des pages s'accroît pour inclure l'information archivée, les résumés méthodologiques, les images des questionnaires et les hyperliens aux documents en ligne. On peut consulter les pages HTML ainsi produites par l'entremise de CANSIM, Statistiques canadiennes et méthodes statistiques, sur le site Web de Statistique Canada.

29. Autres utilisations passives prévues :

- Inventaire en ligne des questionnaires à partir de la BMDI, prévu pour l'exercice 2002-2003.
- Définitions des éléments de données de la BMDI accessibles par les écrans de recherche du CANSIM à la fin de l'exercice 2001-2002. La fonction est actuellement en voie de développement à l'aide d'une solution basée sur CGI-PERL 5.0 et XML.
- Production en ligne des manuels de classification par la BMDI, prévue au cours de l'exercice 2002-2003.
- Intégration du Registre des logiciels de Statistique Canada et de la BMDI.

¹⁵ Application déployée au départ sous forme d'appliquette, cependant, à cause du temps de chargement, elle est déployée comme une application du serveur.

¹⁶ Visual Age 3.5 d'IBM (JDK 1.2, Swing 1.1) utilise Persistence Builder Data pour confectionner des classes d'accès aux données dans cette version. La mise à niveau à cette version nécessite une refonte des classes d'accès aux données parce que Data Access Builder a été supprimé dans la version 3.5 de Visual Age. IBM n'a pas fourni d'outils de conversion aux utilisateurs de Data Access Builder afin de leur permettre de passer de la version 1.1.7 à 1.2 de JDK. Data Access Builder était le seul outil disponible dans Visual Age pour confectionner les classes d'accès à la base de données au moment où l'on a amorcé l'élaboration des écrans de saisie de la BMDI. L'équipe de développement est sur le point de produire une classe génératrice de code qui servira à confectionner des classes d'accès aux données directement par JDBC. Il faudra refaire les classes d'accès pour toute l'application avant d'effectuer la mise à niveau.

¹⁷ La version Internet renferme des renseignements destinés au public. La version intranet présente des renseignements destinés à un usage interne.

¹⁸ La BMDI renferme des renseignements sur les cycles d'enquête précédents; toutefois, cette information sera accessible lorsqu'on aura mis en œuvre une fonction de recherche directe dans la base de données pour consulter de l'information. La fonction de recherche existante du système de saisie sert spécifiquement aux mises à jour.

B. Utilisation active

30. On prévoit mettre au point, pendant l'exercice 2002-2003, le CANSIM Transactions Formatter, qui chargera des données dans CANSIM et des métadonnées dans la BMDI. Ainsi, les utilisateurs de l'Administrateur d'extraction de l'information et de métadonnées (AEIM)¹⁹ pourront coupler les métadonnées de la BMDI à celles du CANSIM.

31. Autres utilisations actives prévues des *métadonnées* :

- Services de codage par une intégration au Système de codage pour la classification des industries (SCCI).
- Édition intelligente des publications électroniques grâce aux *métadonnées* et aux données de diverses sources.

C. Autres facteurs de diffusion

32. Une étude architecturale a eu lieu en 2001²⁰ afin de dégager les interactions entre les principales bases de données de référence de Statistique Canada²¹ et de proposer des technologies qui appuieraient la prestation des services, l'échange des données et le traitement coopératif. Cette étude doit servir de guide aux efforts d'élaboration futurs à long terme à Statistique Canada.

33. L'initiative du Gouvernement en direct (GeD) est l'engagement du gouvernement fédéral à fournir aux Canadiens un accès en ligne à tous les renseignements fédéraux et services fédéraux d'ici l'an 2004 par l'entremise du site Web GeD Canada (canada.ca). La BMDI représentera une source d'information clé pour cette initiative.

34. La technologie évoluant rapidement, les utilisateurs et les producteurs de données font face au défi de migrer vers de nouvelles technologies afin de rendre leurs renseignements et leurs données plus accessibles. Le défi consiste à choisir la ou les technologies qui offrent une longévité maximale en contrepartie du temps et des connaissances qu'il faut acquérir pour appuyer la nouvelle technologie. Le Comité de la stratégie en matière de logiciels et le Centre de développement du Web sont des secteurs clés qui peuvent influencer sur la sélection des outils logiciels disponibles aux utilisateurs à Statistique Canada. Grâce à un jeu harmonisé d'outils logiciels, les producteurs et les utilisateurs de données seraient en mesure de gérer plus efficacement leurs renseignements et données et on fournirait un accès uniforme à des technologies appuyées.

VI. CONCLUSION

35. Les efforts déployés aux fins de la base de métadonnées intégrées au cours des trois dernières années ont porté sur l'élaboration de *métadonnées* et leur chargement dans un entrepôt central, appelé Base de métadonnées intégrées (BMDI). On a utilisé les voies actuelles de diffusion pour fournir aux utilisateurs un accès aux données de la BMDI. L'organisme s'est employé à recueillir les métadonnées plutôt qu'à trouver de nouvelles voies d'accès étant donné que les métadonnées sous-jacentes doivent exister avant que les voies d'accès aient une quelconque utilité.

¹⁹ L'AEIM sert à réunir des microdonnées, des données agrégées et des métadonnées de diverses sources afin de confectionner des fonds de produits statistiques et de consulter des données et des métadonnées à des fins d'analyse ou d'exportation.

²⁰ Hutton, T. et Graves, R. (2001), « Coherence of Reference Databases Architecture Study », Division du développement de systèmes, Statistique Canada, version finale, 11 juin 2001

²¹ Base de métadonnées intégrées (BMDI), Infrastructure de données spatiales, Registre des entreprises, Registre des adresses et microdonnées fiscales.

36. Les activités de collecte et de mise à jour représentent des tâches gigantesques qui exigent des efforts et un engagement substantiels de la part de l'organisme. On a utilisé les technologies orientées objet, l'intranet et XML pour mener à bien ces activités afin de simplifier les tâches liées au projet BMDI. Dans l'ensemble, l'expérience a été positive et stimulante pour le projet BMDI.

37. Le défi consiste maintenant à tenir à jour la BMDI, tant au niveau des données qu'elle renferme que des technologies qu'elle utilise (comme nous avons adopté la technologie orientée objet, le défi est d'éviter l'obsolescence au fil de l'évolution rapide des technologies orientées objet). Au nombre des orientations à venir, notons le passage d'une collecte passive à une collecte plus active et l'utilisation de *métadonnées* de même que de meilleures voies d'accès pour les utilisateurs internes et externes aux *métadonnées* stockées dans la BMDI.

BIBLIOGRAPHIE

Carnahan, L., Gallagher, L. (2000), "A General Purpose Registry/Repository Information Model", Draft, National Institute of Standards and Technology, Information Technology Laboratory.

Ehrenström, B. (2000), "Neuchâtel Group, The Neuchâtel Terminology: Classification Database Object Types and their Attributes", UN/ECE Metis Work Session, Washington D.C. United States, November 28-30, 2000.

Graves, R., Hutton, T., (2001), "Coherence of Reference Databases Architecture Study", Final Version, Statistics Canada, Systems Development Division.

Hutton, T. (1999), "An Overview of Business Survey Processing", Draft, Statistics Canada, Systems Development Division.

ISO/IEC 11179-1, "Specification and Standardization of Data Elements, Part 1, Framework for the Specification and Standardization of Data Elements", Committee Draft, October 1997.

ISO/IEC CD 11179-3, "Information Technology – Data Management and Interchange – Metadata Registries (MDR) – Part 3: Registry Metamodel (MDR3)", June 6, 2001-12-11.

ISO/IEC PDTR 20943-3, "Information Technology – Achieving metadata registry (MDR) content consistency – Part 3: Value domains", February 26, 2001-12-11.

Johanis, P., (2000), "Statistics Canada's Integrated Meta Database Current Status and Future Plans", UN/ECE Metis Work Session, Washington D.C. United States, November 28-30, 2000.

Smiderle, G., (2000), "Storing UES Metadata on the IMDB", Preliminary Report, Statistics Canada, Unified Enterprise Statistics Program.

Systems Development Division, (2001), "Pilot Project to Interface IMDB with the Cansim II Transaction Formatter", Specifications, Statistics Canada.

Statistics Canada, (2000) "Policy on Informing Users of Data Quality and Methodology".

Statistics Canada, (2001), "Policy on Standards".

Statistics Canada, (1998), "Statistics Canada Quality Guidelines", Third Edition.

Statistics Canada, (2001), “The Government On-Line Strategy and Statistics Canada”. Statistics Canada’s Public Report to Treasury Board on Government On-Line.

Systems Development Division, (2001), “Pilot Project to Interface IMDB with the Cansim II Transaction Formatter”, Specifications, Statistics Canada.