

Distr.
GÉNÉRALE

CES/AC.71/2004/11 (Summary)
27 février 2004

FRANÇAIS
Original: ANGLAIS

COMMISSION DE STATISTIQUE et
COMMISSION ÉCONOMIQUE
POUR L'EUROPE

COMMISSION EUROPÉENNE

CONFÉRENCE DES STATISTICIENS
EUROPÉENS

OFFICE STATISTIQUE DES
COMMUNAUTÉS EUROPÉENNES
(EUROSTAT)

**ORGANISATION DE COOPÉRATION ET DE
DÉVELOPPEMENT ÉCONOMIQUES (OCDE)
DIRECTION DES STATISTIQUES**

Réunion CEE/Eurostat/OCDE sur la gestion des systèmes d'information statistique
(Genève, 17-19 mai 2004)

Thème ii): La mise au point de stratégies en matière de technologies de l'information dans les services de statistique

**SYSTÈME INTÉGRÉ DE GESTION DES DONNÉES STATISTIQUES
S'ARTICULANT AUTOUR DES MÉTADONNÉES**

Communication sollicitée
du Bureau central de statistique de la Lettonie¹

Résumé

1. Le présent document a pour objet de décrire l'expérience acquise en ce qui concerne la mise au point et l'application d'un système de traitement des données statistiques de la nouvelle génération, lequel incorpore plusieurs sous-systèmes, et s'articule autour des métadonnées.
2. Il s'agit en l'occurrence, conformément à la stratégie en matière de technologies de l'information du Bureau central de statistique, d'un système centralisé où toutes les données sont stockées dans un entrepôt de données d'entreprise. La nouvelle démarche consiste à regrouper les éléments qui, en toute logique, doivent aller ensemble en faisant appel à des instruments

¹ Document établi par Karlis Zeila (karlis.zeila@csb.gov.lv).

informatiques perfectionnés pour assurer la rationalisation, la normalisation et la coordination des processus de production de données statistiques. Dans la pratique, cela signifie que l'on abandonne l'ancienne approche fondée sur un cloisonnement au profit d'un nouveau mode de traitement des données orienté vers les processus.

3. On a utilisé comme base théorique de l'architecture du système les résultats d'études sur les métadonnées effectuées et publiées par M. Bo Sundgren, Statistics Sweden. Un système de traitement des données s'articulant autour des métadonnées repose sur la systématisation et le stockage, dans une base centralisée de métadonnées, de toutes les métadonnées relatives aux enquêtes, aux indicateurs et aux classifications en vigueur, ce qui permet de réunir les conditions préalables pour une unification et une harmonisation des indicateurs statistiques, tout en évitant les doubles emplois entre différentes enquêtes. L'introduction d'un système unifié d'indicateurs statistiques se traduit par un élargissement des fonctionnalités du système et les métadonnées deviennent l'élément clef de l'ensemble du système.

4. L'intégration du système s'appuie sur l'utilisation directe de données extraites des registres des entreprises pour le traitement des données issues d'enquêtes statistiques auprès des entreprises. Le module «registres» du système prévoit l'emploi de plusieurs registres statistiques différents. La plupart des modules logiciels du système sont reliés au module «registres».

5. Tous les modules logiciels du système sont rattachés au module «métadonnées de base». Nous pouvons dire qu'il s'agit d'un système intégré parce que l'ensemble des données est stocké dans l'entrepôt de données d'entreprise.

6. Le système se divise en une série de modules logiciels d'application de gestion, énumérés ci-après, qui doivent couvrir et étayer l'exécution de toutes les phases du traitement des données statistiques:

- Le module fondamental «métadonnées de base»;
- Le module «registres»;
- Le module «saisie et validation des données»;
- Le module «saisie de données de masse»;
- Le module «collecte de données fondée sur le Web»;
- Le module «imputation de données manquantes»;
- Le module «agrégation de données»;
- Le module «analyse de données»;
- Le module «diffusion de données»;
- Le module «gestion des utilisateurs».

7. Le module «collecte de données fondée sur le Web» s'inscrit dans le prolongement de la méthode de collecte de données au moyen de questionnaires sur papier. La présentation du formulaire électronique est pratiquement la même que celle de la version papier, de sorte que les enquêtés ne sont pas dépayés lorsqu'ils passent de la version papier à la version écran. En fait, la communication électronique des données présente des avantages supplémentaires pour l'enquêté, en ce sens qu'il peut visualiser les données qu'il a fournies pour les périodes précédentes, qu'il a la possibilité de modifier ces informations, le cas échéant, etc.

8. Les principales conclusions que nous pouvons dégager de la mise au point et l'application du système sont les suivantes:

- La conception d'un nouveau système d'information doit s'appuyer sur les résultats d'une analyse approfondie des processus et des flux de données statistiques;
- Il faut que des objectifs clairs, en termes de ce que l'on veut accomplir, soient fixés, examinés et approuvés par toutes les parties intéressées: statisticiens, informaticiens et personnel administratif;
- Grâce à l'étude de faisabilité, nous avons clairement compris que toutes les étapes du traitement des données statistiques des différentes enquêtes peuvent être normalisées, tandis que chaque enquête peut, par ailleurs, exiger une fonction complémentaire (procédures non normalisées) nécessaire pour ce traitement particulier de données d'enquête;
- Pour résoudre les problèmes que posent les interfaces de procédures non normalisées du point de vue de l'exportation ou de l'importation de données, il convient de mettre au point un système qui permette d'utiliser des logiciels standard de traitement des données statistiques ainsi que d'autres logiciels généraux disponibles sur le marché;
- Les deux parties, à savoir les statisticiens et les informaticiens, devraient être associées dès le départ à la conception et l'application d'un système d'information statistique intégré s'articulant autour des métadonnées;
- Une division précise des tâches et des responsabilités entre les statisticiens et le personnel informatique est la clef du succès;
- Les deux parties doivent avoir une idée claire de l'ensemble des processus statistiques, qui seront pris en compte par le système, ainsi que de la signification et du rôle des métadonnées à l'intérieur du système, du point de vue à la fois des producteurs et des utilisateurs de données;
- C'est aux statisticiens et non pas au personnel informatique ou administratif qu'il appartient de prendre l'initiative de renoncer à un système classique cloisonné de production de statistiques pour passer à un système orienté vers les processus;
- Il est essentiel, pour pouvoir remplacer l'environnement de traitement des données existant par un environnement nouveau, que les statisticiens aient la motivation requise à cette fin;

- L'amélioration de la connaissance des métadonnées est l'une des tâches les plus importantes à accomplir tout au long du processus de conception et d'exécution du projet;
- Il est nécessaire de constituer et de former un groupe spécial de statisticiens, qui sera chargé de la tenue à jour de la base de métadonnées et responsable de l'exactitude de ces métadonnées;
- Pour que l'ensemble du système fonctionne de manière optimale, il importe de veiller, par une organisation adéquate, à ce que l'exécution des processus statistiques s'effectue dans le bon ordre;
- Aux fins de la gestion et l'entretien du système, on doit disposer des services d'informaticiens qualifiés possédant une bonne connaissance des applications dans les domaines suivants: administration du serveur 2000 SQL de Microsoft, service d'analyse Microsoft, autres instruments Microsoft, produits de la famille PC-AXIS et modèle de données de système.
