

Distr.
GÉNÉRALE

CES/AC.71/2003/7 (Summary)
3 décembre 2002

FRANÇAIS
Original: ANGLAIS

**COMMISSION DE STATISTIQUE et
COMMISSION ÉCONOMIQUE
POUR L'EUROPE (ONU)**

**CONFÉRENCE DES STATISTICIENS
EUROPÉENS**

COMMISSION EUROPÉENNE

**OFFICE STATISTIQUE DES
COMMUNAUTÉS EUROPÉENNES
(EUROSTAT)**

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

Réunion mixte CEE/Eurostat/OCDE sur la gestion des systèmes d'information statistique
(Genève, 17-19 février 2003)

Point II: Impact des mesures et normes techniques sur la qualité des données

**INNOVATIONS DANS LE DOMAINE DU TRAITEMENT
DE L'INFORMATION À STATISTICS NETHERLANDS**

Document sollicité

Émanant de Statistics Netherlands¹

Résumé

1. Il y a deux ans, Statistics Netherlands a profondément modifié sa structure organisationnelle et a défini sa stratégie pour les années à venir. Trois éléments de celle-ci présentaient un intérêt particulier pour son service informatique. En premier lieu, la structure de Statistics Netherlands a été modifiée: anciennement axée sur les domaines statistiques (et caractérisée par des services cloisonnés), elle est maintenant axée sur les opérations. Deuxièmement, l'accent n'a plus été mis sur la réalisation d'enquêtes en vue de la collecte de données, mais sur l'établissement de liens avec des registres administratifs qui contiennent

¹ Établi par Olav ten Bosch (obos@cbs.nl).

de grandes quantités de données. Troisièmement, le processus de production ne privilégie plus la quantité mais la qualité. Dans le document résumé ici, nous décrivons certaines innovations qui ont été adoptées par le service informatique pour soutenir cette nouvelle stratégie. En outre, nous montrons quelle influence ces innovations ont sur la qualité des données au cours des opérations de production et quels effets elles ont sur la qualité des données publiées.

2. L'une des innovations que nous souhaitons mentionner ici est l'utilisation croissante de la *technologie des entrepôts de données* au cours des opérations statistiques. Nous avons constaté que les entrepôts de données conviennent pour le stockage de grandes quantités de données provenant de registres. Grâce à eux, les statisticiens ont facilement accès à ces données en utilisant des outils normalisés pour le traitement analytique en ligne. Cela peut sembler simple, mais en réalité il n'en est rien. La conception, le chargement et l'entretien d'un tel entrepôt de données représentent une tâche complexe. Cependant, une fois que les données sont chargées et que des cubes sont définis, les statisticiens peuvent mieux explorer les données et le faire avec plus de souplesse que par le passé. En fait, ils ont vu des choses qu'ils n'avaient jamais vues auparavant. On ne sait pas encore si cela influe sur la qualité, mais à mesure que les statisticiens acquièrent des capacités plus puissantes pour analyser leurs données, nous pensons que leurs résultats seront plus complets, ce qui est l'une des dimensions de la qualité des données.

3. Une autre innovation digne d'intérêt est la création de *dépôts de métadonnées*. Ces derniers contiennent des informations normalisées telles que des classifications statistiques, des variables et des mappages. Bien que les dépôts de métadonnées existent déjà depuis longtemps, la nouvelle structure, axée sur les opérations, adoptée par Statistics Netherlands permet mieux de les utiliser effectivement dans le cadre de nouveaux systèmes. En outre, ils ont été ouverts plus largement à l'aide d'interfaces de composants, de façon à pouvoir être utilisés facilement à partir de n'importe quel système d'information de Statistics Netherlands. Nous avons constaté qu'une politique très stricte concernant les métadonnées (obligation d'utiliser les dépôts de métadonnées) ne donnait pas les résultats escomptés. En revanche, il est apparu qu'une politique plus souple en la matière (il faut les utiliser chaque fois que possible) améliore la comparabilité entre les différentes sous-opérations et a, en conséquence, des effets positifs sur la cohérence globale des statistiques.

4. Récemment, Statistics Netherlands s'est mis à utiliser le *RUP* (Rational Unified Process) dans le service informatique. Bien que l'on en soit encore à un stade préliminaire, deux projets pilotes ont déjà été réalisés à l'aide de cette méthode de développement de logiciels. En particulier, l'utilisation d'un cadre prédéfini de définitions des besoins en matière de systèmes d'information s'est avérée extrêmement utile, tant pour les utilisateurs finals que pour les analystes de systèmes. En outre, les systèmes d'information qui ont été conçus et réalisés dans le cadre de ces projets pilotes comportaient très peu d'erreurs de programmation, ce qui indique leur niveau de qualité. D'une manière générale, nous pensons que l'adoption de cette méthode a des effets positifs sur la qualité du développement et de la maintenance du système d'information et, par conséquent, sur la qualité du processus statistique dans son ensemble.

5. La dernière question que nous abordons dans ce document concerne la base de données *StatLine*. En une dizaine d'années, ce système, qui était à l'origine un programme très simple d'établissement de tableaux, est devenu une base de données multidimensionnelle souple, ouverte et aisément adaptable aux besoins, qui permet aux utilisateurs finals d'examiner de façon interactive ses contenus via Internet. Actuellement, StatLine contient *tous* les résultats

statistiques publiés par Statistics Netherlands. StatLine présente l'avantage considérable de permettre un accès à tout sous-ensemble de données via un simple hyperlien concis. Cette possibilité est utilisée depuis environ deux ans et il est apparu qu'elle était extrêmement utile. Dans un avenir rapproché, on élargira ce mécanisme de façon à pouvoir utiliser des technologies Web courantes, ce qui le transformera en *service statistique sur le Web*. Ce dernier offrira aux systèmes externes un accès automatisé au contenu statistique via Internet. Les clients peuvent utiliser ce service pour prendre connaissance de certains indicateurs statistiques à des intervalles réguliers. Une autre amélioration résulte de l'incorporation de technologies SIG spécialisées dans StatLine. Grâce à ces technologies, les utilisateurs peuvent parcourir la base de données via une entrée géographique. Tant le service statistique sur le Web que le système d'entrée géographique améliorent l'accessibilité et la clarté des données et contribuent par conséquent à la qualité des données statistiques publiées.
