

Distr.
GÉNÉRALE

CES/AC.71/2005/26 (Summary)
3 février 2005

FRANÇAIS
Original: ANGLAIS

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

COMMISSION EUROPÉENNE

**CONFÉRENCE DES STATISTICIENS
EUROPÉENS**

**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
(Bratislava, Slovaquie, 18-20 avril 2005)

Thème iv): Examen et suivi des activités de la Conférence des statisticiens européens

**SYSTÈME DE DIFFUSION (STATWEB 4.0) DE DONNÉES ET MÉTADONNÉES
STATISTIQUES À STATISTICS NETHERLANDS**

Document d'appui

Émanant de Statistics Netherlands¹

Résumé

I. INTRODUCTION

1. Le présent document expose les efforts entrepris par Statistics Netherlands (SN) pour rendre plus efficace et plus cohérent le processus de publication des données et métadonnées statistiques. La coordination de l'ensemble des métadonnées utilisées dans le cadre de l'activité de production de statistiques constitue un objectif important, d'autant que Statistics Netherlands ambitionne d'être un institut de normalisation.

¹ Établi par Albert Jacques (ajcs@cbs.nl).

II. COORDINATION

2. Pour arriver à assurer la coordination, on peut, au lieu d'énoncer des règles, proposer des outils. Les outils proposés devront toutefois agir sur toutes les phases de la production de statistiques. L'intégration aussi bien verticale qu'horizontale de la production de statistiques suppose désormais que l'on coordonne toutes les phases de cette production, l'entrée, le traitement et la sortie. Pour la première étape du processus, Statistics Netherlands s'est inspiré de l'expérience de l'Australian Bureau of Statistics (ABS), et cette première phase a abouti à la mise en place d'un serveur de classification pour la coordination des produits statistiques.

3. Les résultats ont été encourageants, mais, en raison des critères d'acceptation, de nombreuses variantes des normes ont dû être acceptées dans le serveur de classification, ce qui en a réduit l'efficacité. En s'appuyant sur le modèle du système Cristal, traduction et extension formalisées des concepts métastatistiques essentiels constituant la structure des bases de données et des tableaux, on pourra faciliter la transformation de données structurées de façon intuitive et insatisfaisante pour obtenir des structures plus rigoureuses et mieux pensées.

III. MÉTADONNÉES

4. À partir du modèle Cristal souple qui avait été pris comme base et des résultats obtenus par le groupe de Neuchâtel, un nouveau type de métadonnées a été élaboré. Les métadonnées se rapportant à la production de statistiques sont groupées en quatre catégories: conceptuelles, opérationnelles, qualitatives et techniques.

5. Les informations (métadonnées) relatives à des statistiques présentées sous forme de tableau, par exemple, peuvent être déterminées même si le tableau en question ne contient pas de résultats statistiques. Dans le modèle néerlandais, ce genre de métadonnées s'appellent «métadonnées conceptuelles». Différents processus peuvent aboutir à des résultats différents, avec les mêmes métadonnées conceptuelles. Pour distinguer entre ces différents résultats, on doit disposer de métadonnées opérationnelles et qualitatives.

IV. NOUVEAU SYSTÈME DE DIFFUSION STATLINE (STATWEB 4.0)

6. Dans le système actuel de diffusion Statline (StatWeb 3), les tableaux sont conçus comme des unités indépendantes. Des variables identiques figurent dans plusieurs tableaux, avec le risque que des données différentes soient publiées pour des indicateurs identiques, et aussi le risque qu'il existe des métadonnées différentes pour des indicateurs identiques apparaissant dans des tableaux distincts. Dans le nouveau modèle StatWeb 4.0, tant l'information statistique que la méta-information peuvent être réutilisées aisément. Il est aussi possible de communiquer des métadonnées statistiques à des utilisateurs extérieurs sans les informations statistiques correspondantes, ce qui s'inscrit dans le droit fil des ambitions de Statistics Netherlands qui aspire à jouer le rôle d'institut de normalisation.

V. NIVEAU DE COORDINATION DES MÉTADONNÉES

7. Toutes les métadonnées expressément désignées comme telles par le Conseil d'administration de SN sont des métadonnées coordonnées. Il peut s'agir soit d'une norme nationale soit d'autres métadonnées coordonnées. Les métadonnées coordonnées sont toujours

partagées. Toutes les autres métadonnées, qui sont donc non coordonnées, figureront néanmoins dans le serveur de métadonnées, éventuellement avec une date de fin de validité. Les métadonnées non coordonnées peuvent être partagées (utilisées pour plusieurs procédures) ou non partagées, auquel cas elles sont réservées à un utilisateur unique.

VI. RÈGLES DE MISE EN ŒUVRE POUR LES MÉTADONNÉES COORDONNÉES

8. Les métadonnées coordonnées définissent des concepts protégés. Ces concepts ne peuvent pas être utilisés par SN dans le même contexte avec une signification différente. Quand les exigences en matière de confidentialité et de qualité le permettent, SN se conforme aux normes nationales pour ses publications. Ces normes nationales prévalent toujours sur les normes coordonnées, qui, elles-mêmes, prennent le pas sur les normes non coordonnées. Lorsqu'on agrège certaines catégories d'une classification, une nouvelle variante de la classification est stockée en même temps que la version originale. Il n'est pas absolument nécessaire de donner des noms à ces variantes, mais on peut le faire le cas échéant. Les métadonnées non coordonnées seront incluses dans le serveur de métadonnées, car les besoins particuliers des utilisateurs d'informations statistiques doivent être pris en compte si Statistics Netherlands le souhaite. Il s'agit souvent de publications supplémentaires s'ajoutant aux statistiques publiées régulièrement qui reposent sur des métadonnées coordonnées. Ces métadonnées non coordonnées sont placées sur une liste de métadonnées «tolérées», valables pour deux ans au maximum. Les métadonnées temporaires, qui n'ont pas encore été coordonnées, sont incluses sur la liste des métadonnées «tolérées», qui indique quelles sont les métadonnées dont on a le droit de se servir uniquement pour les nouveaux tableaux ou pour les mises à jour des tableaux existants. Quant aux anciennes métadonnées, qui ont été utilisées pour des tableaux ne faisant plus l'objet de mises à jour, et dont la date d'expiration est dépassée, on peut les conserver plus longtemps, si besoin est. Les métadonnées non coordonnées peuvent être partagées ou non partagées.

VII. MISE EN PLACE À STATISTICS NETHERLANDS

9. La mise en place combine en fait deux processus. Le premier est l'installation de StatWeb 4.0, et le deuxième l'intégration du serveur C actuel et des autres métaserveurs dans StatWeb 4.0. Les principales innovations dans StatWeb 4.0 par rapport à StatWeb 3 sont les suivantes: a) un système de division des tableaux en petits cubes, et b) un fractionnement des métadonnées, qui sont désormais rattachées aux divers tableaux, des liens renvoyant à un lieu central regroupant toutes les métadonnées. Aux fins de l'introduction de ces innovations, la première étape consiste à opérer une transformation; tous les tableaux de StatWeb 3 doivent être adaptés au modèle fondé sur les petits cubes de StatWeb 4.0. L'unité chargée de la coordination statistique a établi des normes pour la classification par activité, par sexe, par âge, par taille d'entreprise d'après le chiffre d'affaires et par région, qui serviront dans le cadre de la transformation. Pour les autres classifications et concepts, on utilisera les métadonnées centrales actuelles. Pendant les six premiers mois qui suivront la transformation, les tableaux seront établis à la fois en StatWeb 3 et en StatWeb 4.0. Une nouvelle administration sera créée pour les transactions correspondantes. Un plan de communication sera mis sur pied pour informer les utilisateurs au sujet du processus de changement.

VIII. CONCLUSIONS

10. En intégrant les métadonnées dans le système de diffusion Stateline, on rend incontournable leur utilisation, car la publication de statistiques est impossible autrement. Sont comprises parmi les métadonnées non seulement les classifications, mais aussi les définitions et les concepts. Le modèle Cristal constitue un instrument utile pour cette intégration. StatWeb 4.0 est considéré comme une étape importante vers la coordination aux niveaux national et international des informations statistiques produites par Statistics Netherlands. Outre qu'il ambitionne de faire office d'institut de normalisation, Statistics Netherlands espère aussi jouer un rôle important dans la coordination du processus statistique global et faire comprendre l'importance et les avantages de la coordination.
