

Distr.  
GÉNÉRALE

CES/AC.71/2004/20 (Summary)  
8 mars 2004

FRANÇAIS  
Original: ANGLAIS

COMMISSION DE STATISTIQUE et  
COMMISSION ÉCONOMIQUE  
POUR L'EUROPE

COMMISSION EUROPÉENNE

CONFÉRENCE DES STATISTIENS  
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

**LA BASE DE MÉTADONNÉES INTÉGRÉE: UN ÉLÉMENT PRIMORDIAL  
POUR LES TRAVAUX DE DÉVELOPPEMENT**

**Document d'appui**

Présenté par le Bureau central de la statistique de Croatie<sup>1</sup>

**Résumé**

**I. INTRODUCTION**

1. Le présent document a pour objet de décrire le modèle général de base de métadonnées statistiques intégrée et son utilisation pour le traitement des enquêtes statistiques par le Bureau central de la statistique de Croatie (CROSTAT). Le modèle n'est pas encore totalement construit mais il est appliqué en partie.

---

<sup>1</sup> Établi par Zdenko Milonja (zmilonja@dzs.hr).

2. Si nous attribuons aux métadonnées un rôle clef dans les travaux de développement, c'est parce que nous pensons que l'élaboration d'un système statistique ne fait qu'un avec celle des métadonnées. Si elles existent ou si elles sont organisées sous la forme d'une base de données, elles pourraient constituer un bon point de départ pour l'élaboration de tout le système statistique. La centralisation des normes, classifications, méthodologies et autres connaissances concernant la question permet d'établir le système statistique plus rapidement et de le mettre immédiatement à la disposition du public.

## **II. ARCHITECTURE GÉNÉRALE**

### **A. Éléments principaux**

3. L'architecture générale de la base de métadonnées intégrée comprend plusieurs éléments principaux: description de la structure organisationnelle, plan annuel et calendrier de publication, descriptions des enquêtes et classifications, notamment. Outre la base de métadonnées, il existe une base de données statistiques; elle se présente en deux versions: celle des microdonnées et celle des macrodonnées. Le troisième élément est le système de production de statistiques, qui établit un lien entre les données statistiques et les métadonnées statistiques.

4. Les relations entre tous ces éléments sont très complexes dans la pratique mais il est généralement possible de définir le flux de données et la séquence des événements, ce qui revient à produire les données statistiques.

5. Les technologies de l'information jouent un rôle de plus en plus important dans le traitement des enquêtes statistiques et la diffusion de leurs résultats, et elles absorbent des ressources considérables. La base de métadonnées intégrée donne la possibilité d'accroître la productivité à divers niveaux de l'application des technologies de l'information. Dès lors qu'il est possible de séparer les données, mais aussi les métadonnées, des programmes, ceux-ci deviendront probablement plus faciles à adapter.

### **B. Relations fondamentales**

6. La logique élémentaire veut que les métadonnées statistiques soient antérieures aux données statistiques. Pour qu'il soit possible de produire les données statistiques, il faut que les métadonnées existent déjà. Elles se trouvent généralement à l'état latent dans les documents publiés, parmi les connaissances acquises par les spécialistes, etc. Sans métadonnées statistiques, il est impossible de produire des données statistiques et, en ce sens, les métadonnées précèdent les données statistiques. C'est la production des données statistiques qui fait la jonction entre les métadonnées et les données.

## **III. FONCTIONS DE BASE DU SYSTÈME**

### **A. Création et tenue à jour de la structure organisationnelle**

7. La structure organisationnelle est la métadonnée de base en ce qui concerne la planification et la gestion. Du fait de sa complexité, la production fait intervenir un grand nombre de personnes et comprend de multiples activités. Pour la gérer avec efficacité, nous avons fait

en sorte que la base de métadonnées comprenne des données relatives à la structure interne du bureau.

## **B. Production de registres et de listes des personnes interrogées**

8. Une question intéressante se pose concernant la place des registres. Font-ils partie des données statistiques ou des métadonnées statistiques? Nous les mettons dans la catégorie des métadonnées, de même que les listes des personnes interrogées, qui en sont dans l'ensemble dérivées. Les connaissances nécessaires pour établir le registre (les listes des personnes interrogées) sont de l'ordre des métadonnées, alors que les données recueillies au sujet des unités statistiques sont des données statistiques.

## **C. Tenue à jour des classifications**

9. Les classifications font partie de la base de métadonnées et les opérations de traitement et de mise à jour sont effectuées au même endroit et relèvent d'une autorité unique. Nous aimerions mettre au point un système centralisé unique et cohérent pour les classifications. Il est nécessaire pour chaque enquête d'utiliser un système de classification.

## **D. Planification**

10. La production des données statistiques est une opération extrêmement complexe et onéreuse. La planification consiste entre autres à attribuer des tâches aux divers éléments de la structure organisationnelle. Les métadonnées contiennent par conséquent des détails sur l'organisation proprement dite, les tâches à exécuter et le dispositif de planification pour établir le calendrier des tâches à exécuter. Il est également possible de suivre la durée effective de toutes les activités réalisées au cours de la production.

## **E. Création de logiciels**

11. Les technologies de l'information jouent un rôle important dans la collecte, le traitement et la diffusion des données statistiques. Il existe deux méthodes possibles: soit utiliser un logiciel polyvalent, qui est peu coûteux et n'est pas conçu pour répondre à un besoin particulier, soit créer un logiciel adapté à un besoin particulier, ce qui est coûteux. Dès lors que le stockage des métadonnées est centralisé, il est intéressant d'envisager la création automatisée de logiciels à des fins statistiques.

## **IV. ÉTAT D'AVANCEMENT DU PROJET**

12. Les projets complexes et de grande envergure ne peuvent être réalisés très rapidement et de brèves informations seront données sur l'état d'avancement du projet.

-----