

NATIONS UNIES

COMMISSION ECONOMIQUE
POUR L'EUROPE

ОБЪЕДИНЕННЫЕ НАЦИИ

ЭКОНОМИЧЕСКАЯ КОМИССИЯ
ДЛЯ ЕВРОПЫ

UNITED NATIONS

ECONOMIC COMMISSION
FOR EUROPE

SEMINAIRE

СЕМИНАР

SEMINAR

COMMISSION DE STATISTIQUE et
COMMISSION ÉCONOMIQUE POUR L'EUROPE



CONFÉRENCE DES STATISTICIENS EUROPÉENS

CER/SEM.35/SV/R.1(SUM)
26 avril 1996

Original: FRANCAIS

Séminaire sur les Statistiques Officielles - Passé et Futur
(Lisbonne, Portugal, 25-27 septembre 1996)

SESSION V: NOTRE LEGAT POUR LES GENERATIONS A VENIR

**DOCUMENTATION ELECTRONIQUE DES OUTILS STATISTIQUES:
UNE EXPERIENCE FRANCO-ROUMAINE**

Rapport présenté par l'Institut national de la statistique et
des études économiques (INSEE) (France) et
la Commission Nationale de statistiques (Roumanie) 1/

Resumé

LES FONDEMENTS

Historique

1. La préoccupation de l'INSEE pour la documentation de ses outils statistiques est née il y a une quinzaine d'années, dans un double souci de transmission du savoir-faire à l'intérieur de l'Institut et de recherche de cohérence dans le traitement de données. A l'origine, c'était principalement le point de vue des producteurs qui était privilégié, mais cette situation tend aujourd'hui à évoluer avec celle de la demande qui s'exprime en matière de données statistiques.

2. Cette préoccupation s'est traduite dans une méthode et un outil informatique. La méthode porte le nom de DDS (Dictionnaire de données Statistiques), l'outil informatique la traduit en termes opérationnels, sous trois formes à l'heure actuelle.

1/ Préparé par Dominique Crosnier, (INSEE) France et Liana Marina, CNS, Roumanie.

GE.96-

Les principes de la méthode

3. Elle s'appuie sur une modélisation des métadonnées à l'aide d'objets qui permettent de décrire et classer les différents éléments textuels (notes, textes de programme,...) concourant à la mise en oeuvre d'une opération statistique dans ses différents phases, et sur la définition des liens que ces différents éléments entretiennent entre eux. Ainsi, tout texte contenant de l'information utile à la compréhension d'une phase de l'opération statistique sera classé sous forme d'une fiche constituée de trois parties: la description, par ses propriétés, de l'objet auquel correspond le texte; le texte lui-même, les liens qui existent entre ce texte et d'autres qui sont classés dans d'autres fiches.

L'outil informatique

4. Il permet, entre autres, de stocker, mettre à jour, rechercher, éditer ces textes de référence. Il existe sous trois formes qui ont en commun d'utiliser le même SGBD, en l'occurrence ORACLE: L'une a reçu le nom de modèle Enquête; la deuxième de modèle TEXO; la troisième, qui a une vocation plus transversale, est destinée à constituer un inventaire des outils utilisés par le Système Statistique Public français et porte le nom d'ISIS. Dans le modèle Enquête, comme dans ISIS, la modélisation des données est fixée. En revanche, dans le modèle TEXO, une partie de la modélisation reste à la discrétion de l'utilisateur. Ces trois versions existent sous MVS, l'éditeur de texte étant ISPF. Ces développements ont été réalisés en 1990, avec la version 5 d'ORACLE.

5. A la suite de cette première phase, le modèle Enquête a connu deux développements. D'abord une version micro monoposte avec Oracle V.5., l'éditeur de texte étant Word 5.5; puis une version réseau local sous UNIX. Cette dernière a été développée en 1993-1994 avec la CNS de Roumanie, dans le cadre d'actions de coopération entre les deux Instituts, financées par le programme PHARE.

LA PRATIQUE

Les réalisations

6. A l'INSEE, la documentation des outils relève du volontariat. Elle n'est assortie d'aucune obligation institutionnelle. A ce jour, il existe 25 bases qui contiennent en moyenne 1500 fiches. Ce sont principalement des outils statistiques concernant les ménages qui ont été documentés. (Emploi, Logement, formation et qualification professionnelle...). Pour des enquêtes réalisées en partenariat et pour lesquelles l'INSEE s'était engagé à faire la

documentation c'est cette méthode et cet outil qui ont été utilisés systématiquement (Transport, Santé, Mobilité et Insertion Sociale). Depuis la création de la version micro, d'autres composantes du Système Statistique Public utilisent cet utile.

7. En Roumanie, après la période d'apprentissage des années 1992-1994, la CNS a commencé à utiliser ces outils. La première expérience, jugée satisfaisante porte sur l'enquête Emploi roumaine (AMIGO).

Les difficultés

8. Les obstacles au développement de ces procédures tiennent pour une bonne part aux caractéristiques de l'outil: manque de souplesse et d'adaptation aux techniques informatiques récentes, la méthode elle-même étant largement appréciée. Cette situation a conduit à définir une évolution du projet, de manière à faciliter les recours toujours souhaité à cette procédure.

LES PERSPECTIVES

9. A la lumière de l'expérience accumulée depuis quelques années, une rénovation de l'outil informatique a été entreprise. Elle se fait en collaboration avec la CNS de Roumanie.

Les principes de la rénovation

10. Les lignes générales en sont les suivantes:

- amélioration de l'ergonomie par recours aux interfaces graphiques;
- amélioration de la souplesse: un seul outil permet de générer l'application de modélisation des données voulu selon les cas traité;
- amélioration de la portabilité: offrir la possibilité de pouvoir utiliser les outils avec comme support des données, d'autres moteurs de gestionnaires de bases de données (Paradox, Sybase...);
- amélioration de l'aide aux diffuseurs: offrir la possibilité en sortie d'une base documentaire de récupérer des définitions utilisables avec des liens hypertexte en accompagnement de données diffusées.

Les techniques utilisées

11. Pour atteindre les objectifs fixés, cette rénovation doit faire appel principalement à deux techniques récentes: la programmation orientée objet (ergonomie, souplesse) et l'architecture client/serveur (portabilité).
