

Distr.
GÉNÉRALE

CES/AC.71/2003/24 (Summary)
6 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 I: Mesures à prendre pour améliorer la qualité au niveau de la gestion de la TI

FUTURE ARCHITECTURE INFORMATIQUE

Document d'appui

émanant du Bureau central de statistique d'Irlande ¹

Résumé

Objectifs stratégiques du Bureau central de statistique (*Central Statistics Office - CSO*) d'Irlande relatifs à la technologie de l'information

1. Le CSO a exposé, dans l'énoncé de sa stratégie 2001-2003 (que l'on trouvera sur son site Internet), les objectifs de sa stratégie relative à la technologie de l'information. Les principaux de ces objectifs sont les suivants:

- Faire migrer les systèmes déjà anciens existants du CSO, basés sur des processeurs DEC Alpha, vers un environnement client-serveur, et les reconfigurer;

¹ Établi par Tom Keane (tom.keane@cso.ie).

- Mettre en application les systèmes reconfigurés dans le cadre des paramètres du modèle de données d'entreprise du CSO et des bases de données qui y sont associées (base(s) de données d'entrée, base(s) de données d'archives unitaires épurées, base(s) de données agrégée(s) et base de données de diffusion, et bases de métadonnées), et élaborer les applications nécessaires pour les appuyer;
- Regrouper la centaine d'enquêtes du CSO en six grandes catégories correspondant aux procédures internes des entreprises;
- Mettre en place une couche application utilisateurs communs comportant des fonctions communes à toutes ces catégories et des couches application utilisateur pour des fonctions additionnelles propres à chaque catégorie;
- Faire en sorte de pouvoir utiliser le Web pour la réception et la diffusion de l'information;
- Établir une interface entre les solutions proposées et certains systèmes client-serveur et bases de données Sybase existants, par exemple le Registre central des entreprises (CBR) et le Système de classification et de normes s'y rapportant (CARS).

Ces objectifs, de même que la migration des stocks de données existants vers le nouvel environnement, doivent être réalisés pour la fin de 2005.

2. L'énoncé de la stratégie dit ceci:

«... À mesure que les systèmes centraux migreront vers le nouvel environnement client-serveur, on procédera à la reconfiguration des systèmes selon une démarche générique, à savoir que l'on créera la fonctionnalité commune nécessaire aux différents systèmes d'enquête dans un modèle central, qui sera ensuite copié et adapté au système de traitement propre à chaque enquête...».

et

«... L'usage des technologies électroniques doit être élargi à la collecte et la diffusion des données statistiques...».

Future architecture informatique

3. La future architecture informatique comprendra les architectures des applications, des données et des technologies nécessaires pour concrétiser la vision stratégique de la technologie de l'information du CSO. Ses principaux éléments sont brièvement évoqués ci-après.

4. Le modèle d'architecture informatique de base dont a besoin le CSO consiste en un entrepôt de données avec, en amont et en aval, des applications multicanaux permettant de collecter et de diffuser les données auprès de sources électroniques diverses ainsi que des sources plus traditionnelles, telles que les supports papier.

5. Entre les couches multicanaux, on trouvera des suites d'applications expressément conçues pour la «gestion» de toutes les enquêtes. Le terme de «gestion» englobe la manipulation et la

mise en forme des données provenant des réponses aux enquêtes et de sources administratives pour en faire des produits statistiques susceptibles d'être diffusés par des canaux électroniques multiples.

6. Cette architecture reposera sur une suite d'outils de productivité destinée à permettre le bon fonctionnement du CSO.

7. Il sera tenu compte dans le nouvel environnement des besoins liés à la gestion des données d'entreprise puisque tous les systèmes proposés seront élaborés dans le cadre des paramètres du modèle de données d'entreprise, comme il est expliqué à la section III.2 du document.

8. On trouvera dans le texte intégral du document (disponible en anglais seulement) un schéma de l'architecture globale.

Aperçu des enquêtes et futur cadre de gestion des données du CSO

9. **Historique du fonctionnement du CSO:** Le CSO mène une centaine d'enquêtes portant sur les entreprises et sur les ménages. Leur traitement a été informatisé séparément dans un environnement DECAAlpha VMS selon l'une ou l'autre des méthodes suivantes:

- Le travail de développement des systèmes est effectué en totalité par du personnel local, comme cela est le cas dans les Divisions de l'agriculture et du commerce extérieur, généralement en SAS (www.sas.com) et en TPL (www.qqqsoft.com).
- Le travail de développement des systèmes est effectué en partie ou en totalité à l'échelon central par la Section de développement des systèmes (SDS), généralement en COBOL.

Le personnel local procède aussi à des analyses complémentaires à partir des fichiers de données épurés provenant des systèmes qui ont été développés par la SDS.

10. **Démarche proposée pour la gestion des données dans le cadre du système de gestion des données:** Les systèmes seront reconfigurés dans le cadre des paramètres du modèle de données d'entreprise et des bases de données qui y sont associées (base(s) de données d'entrée, base(s) de données d'archives unitaires épurées, base(s) de données agrégée(s) et base de données de diffusion, dont on trouvera la description ci-après). Par ailleurs, les métadonnées nécessaires pour identifier et expliquer toutes les données statistiques des quatre bases de données précitées seront gérées dans des bases de métadonnées.

11. **Base de données d'entrée:** C'est dans cette base de données (ou dans ce groupe de bases de données) que les données (y compris celles provenant de l'EDI ou de sources administratives) seront initialement introduites après leur saisie, pour être ensuite mises en forme et actualisées.

12. **Base de données d'enregistrements unitaires épurés:** On trouvera dans cette base de données (ou dans ce groupe de bases de données) un instantané des tableaux de la base de données d'entrée tels qu'ils se présentent après un premier traitement. Les tableaux de cette base de données seront statiques et ils porteront une date permettant de les identifier facilement par la suite.

13. **Base de données agrégée:** Au départ, les tableaux de la base de données agrégée (ou du groupe de bases de données agrégées) contiendront les premiers résultats issus de l'agrégation (ou de la tabulation) des données figurant dans la base de données d'archives unitaires épurées. Il faudra souvent soumettre ultérieurement à d'autres traitements (contrôle de la qualité au macroniveau et/ou corrections des variations saisonnières, par exemple) les données de la base agrégée, laquelle sera donc évolutive. La base de données de diffusion contiendra des instantanés de la base de données agrégée, ce qui signifie que les données ne subiront plus aucun traitement une fois qu'elles auront quitté la base de données agrégée.

14. **Base de données de diffusion:** Cette base de données renfermera des données finales pour toutes les périodes disponibles. Les tableaux de la base de données agrégée ne seront pas nécessairement tous introduits dans la base de données de diffusion. Il se peut en effet que certaines des données de la base de données agrégée aient été destinées uniquement à des analyses internes ou à des vérifications. À des fins de contrôle, une date devra être apposée sur les données des tableaux de la base de données de diffusion. Toutes ces données seront statiques.

15. **Base de métadonnées:** La base (ou le groupe de bases) de métadonnées contiendra toutes les métadonnées du système ainsi que les métadonnées nécessaires pour identifier et expliquer toutes les données statistiques figurant dans les bases de données. On y trouvera par exemple une description des variables statistiques, le nom des experts responsables, le fondement juridique des enquêtes, et, dans certains cas, des références ou des liens vers la documentation pertinente.

16. **Autres mémoires-fichiers:** En plus des quatre bases de données et de la base de métadonnées susmentionnées, on tiendra deux mémoires-fichiers qui renfermeront: i) toutes les données transmises au CSO sous forme de fichiers plats, de courriers électroniques, etc.; et ii) toutes les données diffusées non structurées telles que des rapports en format HTML ou PDF, etc.

17. **Circulation des données dans le cadre du système de traitement statistique:** On trouvera dans le texte intégral du document (disponible en anglais seulement) une figure qui permet de suivre le cheminement des données tout au long du cycle de traitement statistique prévu dans le nouveau cadre de gestion des données.
