

Distr.
GÉNÉRALE

CES/AC.71/2001/2
8 mars 2001

FRANÇAIS
Original : ANGLAIS

COMMISSION DE STATISTIQUE et
COMMISSION ÉCONOMIQUE POUR L'EUROPE

CONFÉRENCE DES STATISTICIENS EUROPÉENS

Réunion CEE-Eurostat sur la gestion de la technologie
de l'information en statistique

(Genève, Suisse, 14-16 février 2001)

RAPPORT DE LA RÉUNION DE FÉVRIER 2001

1. La Réunion sur la gestion de la technologie de l'information en statistique s'est tenue à Genève (Suisse) du 14 au 16 février 2001. Y ont participé des représentants des pays suivants : Albanie, Allemagne, Arménie, Bosnie-Herzégovine, Canada, Chypre, Danemark, Estonie, États-Unis, ex-République yougoslave de Macédoine, Finlande, France, Hongrie, Irlande, Israël, Italie, Lettonie, Lituanie, Norvège, Pays-Bas, Pologne, Portugal, République de Moldova, République tchèque, Roumanie, Royaume-Uni, Slovaquie, Slovénie et Suède. La Commission européenne était représentée par Eurostat. Le Japon a également participé à la réunion en application de l'article 11 du mandat de la CEE. Des représentants de l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO), du Fonds monétaire international (FMI), de l'Organisation de coopération et de développement économiques (OCDE), de l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO), de l'Organisation des Nations Unies pour le développement industriel (ONUDI), de l'Organisation mondiale du commerce (OMC) et de la Banque des règlements internationaux (BRI) étaient également présents.

2. L'ordre du jour provisoire a été adopté.

3. M. Mel TURNER (Canada) a été élu Président et MM. Karlis ZEILA (Lettonie) et Enrico GIOVANNINI (OCDE) Vice-Présidents.

4. Les questions de fond ci-après ont été examinées à la Réunion :

a) Défis et possibilités pour les services de statistique travaillant en réseau;

GE.01-30519 (F)

- b) Impact de l'entreposage des données sur la gestion des services de statistique;
- c) Harmonisation des données (d'enquêtes) statistiques avec les données (administratives) issues de registres;
- d) Gestion des ressources dans les services de statistique et rôle des sections informatiques.

5. Les participants ci-après ont joué le rôle d'animateurs : M. Wolfgang KNUEPPEL (Eurostat) pour la question a); M. Lars RAUCH (Suède) pour la question b); M. Christophe ALVISET (France) pour la question c); M. Richard SWARTZ (Census Bureau des États-Unis) pour la question d).

6. Les participants ont recommandé de créer un site Internet qui exposerait les meilleures pratiques pour l'application des technologies de l'information et de la communication en statistique et pour le transfert du savoir-faire dans ce domaine. Le secrétariat a été prié de s'informer des possibilités de mener cette tâche à bien.

7. Les participants ont recommandé qu'une nouvelle réunion sur la gestion de la technologie de l'information soit organisée en février 2003, et que le texte ci-après soit donc inclus dans la présentation intégrée pour 2002/2003 du programme de travail de la Conférence des statisticiens européens :

2.1 Gestion de l'infrastructure informatique

Activités de la CEE

La Réunion sur la gestion de la technologie de l'information en statistique en 2002/2003 examinera les sujets suivants :

- i) Mesures pour améliorer la qualité de la gestion informatique;
- ii) Impact des mesures et normes techniques sur la qualité des données;
- iii) Comment maîtriser efficacement la complexité technique croissante.

8. Les principales conclusions auxquelles les participants sont parvenus au cours de leurs discussions sont présentées dans l'annexe.

ANNEXE

Principales conclusions dégagées lors de la Réunion sur les questions de fond inscrites à l'ordre du jour

A. Défis et possibilités pour les services de statistique travaillant en réseau

1. Comme le développement de la production des statistiques dépend dans une large mesure de la technologie de l'information, les services statistiques font face à des défis très similaires dans ce domaine. Les participants ont examiné la situation actuelle en matière de mise en réseau dans les services de statistique et ont déterminé les domaines où certaines données d'expérience (positives ou négatives) peuvent être mises en commun et les domaines où des problèmes empêchent la poursuite du développement.
2. Le travail en réseau a des effets sur les trois phases de production des statistiques : collecte, traitement et diffusion des données. En outre, certaines questions qui se posent à l'échelle horizontale telles que la gestion, la formation et l'appui ont un caractère intersectoriel. Pour gérer la technologie de l'information, il faut évaluer les risques et les possibilités découlant de l'évolution technologique et la façon dont ceux-ci affectent le processus de production de l'information. Le souhait d'innover de manière accélérée peut conduire à agir prématurément en prêtant insuffisamment attention à des aspects critiques tels que la conception minutieuse des systèmes, la continuité des pratiques existantes, etc.
3. La mise en réseau crée de nombreuses possibilités d'améliorer les travaux des services de statistique. Cependant, de nouveaux problèmes se posent et les gestionnaires doivent les régler : administration et surveillance à distance des systèmes d'exploitation, protection des données et de la vie privée, intégration des données, besoin d'aide en temps réel, besoin de formation aux technologies les plus récentes, augmentation constante des coûts des nouveaux logiciels et matériels informatiques et évolution du rôle de ceux qui gèrent la technologie de l'information.
4. La collecte électronique de données brutes au moyen de technologies faisant appel à Internet est une évolution nouvelle qui doit permettre de réduire les délais et les frais administratifs. Pour ne pas perdre pied face aux instruments qui apparaissent dans le domaine des transactions électroniques, de nouvelles approches sont nécessaires pour la collecte électronique de données. Cependant, le recours à de telles solutions est encore assez peu courant. Les participants ont examiné les facteurs qui entravent l'exploitation de la collecte de données électroniques. Plusieurs pays ont insisté sur la nécessité de disposer de normes et de directives concernant les orientations à suivre en la matière. Des solutions sont aussi nécessaires pour régler des problèmes techniques pratiques tels que celui des questionnaires qui doivent être remplis par de multiples destinataires ou en mode multiséssion.
5. Les entreprises qui répondent aux questionnaires font souvent pression pour pouvoir disposer de moyens électroniques de fourniture des données, mais dans certains pays la législation n'a pas encore atteint un niveau permettant de les utiliser pleinement. En outre, les entreprises commerciales n'ont pas encore compris l'importance de la sécurité dans ce processus et sont prêtes à sacrifier celle-ci pour profiter de solutions meilleur marché ou plus simples. Les mécanismes perfectionnés de sécurité sont souvent jugés trop lourds à gérer et on ne peut

guère s'attendre à ce que leur usage se développe tant que l'on n'aura pas trouvé des solutions conviviales pour l'utilisateur aux problèmes liés aux interfaces technologie-usager.

6. La collecte électronique de données nécessite des compétences et des instruments nouveaux qui entraîneront un accroissement des coûts dans la phase de développement, mais qui pourront déboucher sur une réduction des dépenses à long terme. Les solutions classiques fondées sur EDIFACT fonctionnent correctement mais sont trop onéreuses pour s'appliquer aux entreprises petites et moyennes. Plusieurs services utilisent des questionnaires électroniques pour recueillir des données. La réduction prévisible des coûts de l'Internet permettra de passer d'une formule consistant à remplir et envoyer par courrier électronique des questionnaires prédéfinis à une formule interactive. La Hongrie a décrit un nouveau projet pilote pour la collecte électronique de données via Internet dans des conditions de sécurité. L'emploi du Web pour la collecte de données devrait être accompagné de fonctions d'aide en ligne et le service assuré par le réseau doit être fiable pour permettre le transfert des données requises.

7. Plusieurs services ont fait part de leur expérience pour ce qui est de garantir la sécurité du réseau. Dans de nombreux cas, la technologie disponible (SSL, infrastructure à clefs publiques et réseau privé virtuel (RPV) par exemple) ne semble pas tout à fait au point ou paraît difficile à utiliser. Il est impossible de construire un système avec une sécurité à toute épreuve dans tous les domaines. Il faut dans la pratique accepter un certain degré de risque. Il peut être utile de confier les services de sécurité à des entreprises spécialisées dans ce domaine.

8. Le Web offre d'énormes possibilités pour améliorer la diffusion des produits statistiques. En même temps, il accroît les attentes des usagers en ce qui concerne les délais, la disponibilité, l'accessibilité, l'intelligibilité et l'ergonomie. Le nombre et la diversité des usagers ont fortement augmenté. Parmi ces usagers figurent des experts hautement qualifiés, mais aussi des profanes, des étudiants, etc. Leur compréhension de l'économie, des statistiques et des ordinateurs et leurs besoins en données varient sensiblement. On peut noter que l'importance d'Internet comme moyen de diffusion augmente constamment. Il est probable qu'à l'avenir on jugera souvent la qualité du travail des services statistiques en fonction de la qualité de leurs pages Web. La conception de pages Web fait appel à des techniques similaires à celles qui sont nécessaires pour concevoir une étude et il faut prêter l'attention voulue aux compétences et au personnel requis à cet effet.

9. Des méthodes, des outils et des techniques existent pour évaluer et améliorer la conception des sites et pages Web de diffusion des données. Tout récemment, on s'est beaucoup intéressé à leur ergonomie, c'est-à-dire la mesure dans laquelle un site Web aide l'utilisateur à mener à bien sa tâche spécifique. Pour effectuer les tests d'ergonomie, il faut mesurer les caractéristiques de l'interaction être humain-ordinateur lors de l'utilisation d'une page Web. Les techniques d'ergonomie simplifient les activités finales (maintenance du site, réduction du nombre de demandes d'information et de plaintes, diminution de la charge sur les serveurs et les réseaux), mais accroît la complexité et les coûts de la phase de développement. Les services statistiques peuvent appliquer progressivement les tests d'ergonomie. Un test simple peut être effectué sans technologie particulière et ne prendre que quelques heures sur le temps de travail d'un employé. À l'autre extrême, on peut effectuer des tests exhaustifs.

10. Lorsque des services de collecte et de diffusion de données sont fournis sur Internet, les usagers s'attendent à recevoir une aide en permanence et à obtenir une réponse rapide. En même temps qu'une politique de diffusion, les services de statistique doivent appliquer une politique d'aide aux usagers. L'aide doit être à la fois technique et fonctionnelle. Il existe divers niveaux d'aide (technique ou conceptuelle, par exemple) et il faut clairement indiquer à quel niveau de service on peut s'attendre pour chacun de ces domaines. Certains bureaux de statistique expérimentent des systèmes de réponse automatique à des questions simples. Un aspect important est que la réaction de l'utilisateur permet d'obtenir des informations précieuses pour améliorer l'ensemble du processus de production. En outre, l'accroissement de la diffusion par le Web soulève la question de la nécessité pour les bureaux de statistique de protéger leurs publications électroniques et d'utiliser des techniques pour garantir le respect des droits d'auteur applicables à ces publications.

11. L'emploi d'Internet augmentera fortement et finira peut-être par prendre une place dominante, mais le rythme d'absorption sera lent. Il faut donc continuer à alimenter les canaux existants, parallèlement à l'exploitation de ces nouvelles possibilités, et il est peu probable que les coûts globaux diminuent sensiblement à court terme.

12. L'Internet permet de rendre des données disponibles à partir de n'importe quel endroit, mais le problème pour l'utilisateur est de trouver cet endroit. Il a été suggéré de renforcer l'intégration des données et des métadonnées entre les organisations internationales en plaçant sur les sites Web de celles-ci des liens avec les métadonnées plus détaillées sur les réseaux nationaux. Un éventuel portail unique pour ce type d'accès pourrait être fondé sur le Tableau d'affichage des normes de diffusion des données (TAND) du FMI. Les participants ont été informés d'un projet similaire lancé à la session de travail de novembre 2000 sur les métadonnées statistiques. Les problèmes clés à régler pour garantir le succès de cette approche concernent le langage, l'harmonisation des termes, la variation de structures et les ressources.

13. Le travail en réseau peut donner aux services statistiques la possibilité de mettre en commun des données et métadonnées et d'éliminer les redondances. Il faut pour cela harmoniser les concepts, les classifications, etc. Les organisations internationales devraient jouer un rôle important en définissant des normes dans ce domaine. Sur le plan pratique, l'harmonisation des concepts peut être très difficile, voire même impossible dans certains cas. Il est souvent plus efficace de disposer d'outils pour négocier le mappage des concepts dans des domaines particuliers plutôt que tenter de se mettre d'accord sur des concepts normalisés communs. On peut recommander de suivre au maximum les normes et d'indiquer ensuite les écarts par rapport à celles-ci. Ceci amène à se poser la question de savoir si l'on dispose de définitions des données suffisamment détaillées et de techniques suffisamment efficaces pour recueillir et stocker ces métadonnées. Des solutions reposant sur l'intelligence artificielle pourraient être nécessaires pour saisir le sens des concepts.

14. L'importance de la mise en réseau peut être observée non seulement au niveau technique mais aussi au niveau organisationnel à l'intérieur des pays et entre les pays. Les nouvelles évolutions en matière de technologie de l'information et la nécessité d'en tirer efficacement parti encouragent les pays à accroître la coopération concernant les technologies de l'information statistique. Certains éléments communs favorisent le développement dans le domaine de la technologie de l'information : besoins communs des usagers, préoccupations des fournisseurs de

données (confidentialité), préoccupations administratives (appel à des sources extérieures, utilisation des ressources administratives), coopération entre instituts nationaux de statistique (échanges de données d'expérience, de logiciels et de manuels), intégration européenne. Ceci a poussé Eurostat à jouer un rôle dans l'établissement et l'application d'une politique communautaire sur la technologie de l'information aux fins des statistiques. Des approches communes peuvent être observées pour trouver des solutions concernant la diffusion des données, les préoccupations relatives à la protection de la confidentialité, les méthodes concertées pour combiner les systèmes d'information dans les services statistiques, le recours accru aux sources administratives et la promotion de l'emploi des normes pour les échanges de données. Les travaux de recherche devraient fortement contribuer à régler ces questions.

15. Les nouvelles technologies fondées sur Internet doivent aussi permettre de faciliter le travail en groupe grâce à la création d'extranets sur le réseau public. Divers projets exploitant cette formule ont été présentés à titre d'exemple. Les extranets peuvent fournir des moyens pour la réalisation de travaux à caractère collectif, contribuer au processus de production statistique, aider à coordonner les systèmes statistiques et fournir des informations aux principaux utilisateurs de ces systèmes. On a débattu des futures évolutions envisageables dans ce domaine et des applications qu'elles pourraient avoir dans le système statistique européen. Souvent, il n'y a pas de solutions toutes faites pour utiliser les extranets et ceux-ci doivent être adaptés en fonction de chaque situation. L'une des nouvelles évolutions prometteuses est l'emploi du format XML pour la communication en temps réel de messages avec les partenaires administratifs.

16. La formule extranet est particulièrement utile dans les pays où la répartition du système statistique permet à des personnes travaillant dans les services statistiques de différents ministères de travailler ensemble. Certains pays disposent déjà d'infrastructures de communication relevant de l'administration publique et faisant appel à un extranet ou sont en train d'en établir. Ceci permettra un usage mutuel des sources administratives au sein de l'administration publique et des registres fondamentaux à l'échelle nationale tels que le registre démographique, le registre du commerce, etc.

17. Il a été souvent mentionné dans plusieurs contextes que le format XML était une technologie particulièrement prometteuse pour les organismes statistiques. Il est privilégié en tant que moyen de décrire des structures de données et comme format d'échange. Il doit permettre d'unifier les multiples normes existantes pour assurer l'interface entre des systèmes et des sources de données disparates.

B. Impact de l'entreposage des données sur la gestion des services de statistique

18. La création de bases de données de sortie et d'entrepôts de données peut modifier le paradigme de la production des statistiques, traditionnellement axé sur le domaine d'étude. Les participants ont considéré les conséquences que peut avoir sur le plan de la gestion l'adoption d'une approche intégrée qui remplacerait le système cloisonné actuel.

19. Bien que la création d'entrepôts de données en soit encore à ses débuts et que l'on ne possède pas encore une grande expérience en la matière, certains services de statistique ont donné des exemples du type de changement qui s'est révélé nécessaire au niveau de

l'organisation. Il faut non seulement repenser les modes de production individuels, mais de surcroît amener les statisticiens à une transformation culturelle de leur conception de la production des statistiques. Une stratégie est nécessaire pour diffuser la culture de l'intégration au sein du service et résoudre les problèmes de personnel qui y sont associés. Il peut être utile d'instaurer un roulement du personnel concerné ou de créer une centrale de membres du personnel qui peuvent être détachés auprès des principales sections de statistique. Il faudrait dans le même temps veiller à maintenir les personnels spécialisés dans les domaines d'étude. Parallèlement au système intégré de production, il sera toujours nécessaire de disposer d'équipes spécialisées dans tel ou tel domaine particulier.

20. Contrairement au système cloisonné, la création et la gestion d'entrepôts de données et de bases de données centralisées ne relèvent pas de la compétence d'un département spécialisé dans un domaine d'étude. L'harmonisation et l'unification d'opérations qui se trouvent disséminées au sein de l'organisation seront impossibles sans une gestion solide et efficace. L'administration est un facteur important de réussite et il est nécessaire d'établir une structure élaborée pour administrer l'entrepôt de données, avec une hiérarchie qui remonte jusqu'au niveau de l'organisme de statistique.

21. Plusieurs services ont opté pour la mise en place progressive d'entrepôts de données et de systèmes intégrés d'information statistique. Ils tirent parti des systèmes existants et créent des interfaces avec les nouveaux systèmes parce que la production ne peut être interrompue. Il peut être recommandé d'adopter des solutions qui ont été essayées et qui ont fait leur preuve dans d'autres environnements. Il serait bon de pouvoir mesurer le degré d'intégration de différents flux de production. Cela pourrait contribuer à suivre la démarche adoptée et à mobiliser l'attention et le soutien de la direction et du personnel.

22. Le temps devient un élément décisif lorsque les informations s'accumulent dans l'entrepôt de données. La structure de traitement doit être axée sur un traitement continu. Il faut pour cela abandonner un système caractérisé par des chiffres préliminaires et des chiffres définitifs pour en adopter un autre qui comportera systématiquement une marge d'erreur. La marge d'erreur dont s'assortissent les résultats diminue avec le temps à mesure que l'on dispose de données plus précises. La qualité sera assurée en continu et non à intervalles sporadiques. De la sorte, les utilisateurs seraient libres de choisir entre différents paramètres de la qualité des données, leur actualité ou leur précision par exemple.

23. La préservation de la cohérence des données est l'un des principaux problèmes rencontrés lors de la création d'un entrepôt de données. Il est nécessaire, pour le résoudre, de faire appel à des systèmes de métadonnées perfectionnés et à des classifications harmonisées. Même si les opérations de production utilisent de plus en plus de métadonnées statistiques, il faut encore réaliser leur unification et leur harmonisation au niveau du service. Il faudrait utiliser les métadonnées appropriées pendant tout le cycle de production, dès la collecte des données. C'est la direction qui doit se charger de créer un système de métadonnées dont la gestion est centralisée, en tenant compte des différents intérêts qui existent dans une vaste organisation telle qu'un service national de statistique. Bien souvent, il ne suffit pas de signaler les divergences par rapport aux règles en vigueur; des instruments sont nécessaires pour évaluer les erreurs dues à un manque d'uniformité des définitions. Par ailleurs, il est parfois nécessaire de maintenir différentes séries de données pour être en mesure d'analyser des phénomènes caractéristiques de

la situation à une époque donnée. Une bonne solution consisterait à créer des instruments automatisés à partir d'une description de métadonnées afin que les données concordent les unes avec les autres. Cela dit, l'état actuel des systèmes de gestion des métadonnées ne permet pas encore de mettre au point de tels instruments.

24. Considérant la quantité de données confidentielles stockées dans les entrepôts de données et les systèmes intégrés d'information statistique, il est nécessaire de réexaminer l'environnement de la production du point de vue de la confidentialité. Le développement spectaculaire des possibilités de combinaison des données provenant de sources différentes, les fonctions de recherche, les dispositifs de tabulation à plusieurs variables, etc., obligent à mettre au point des moyens de protection du caractère confidentiel des données. Il faudrait éviter en particulier que tout un chacun puisse utiliser des microdonnées anonymes pour des travaux d'analyse en raison du risque très élevé de divulgation. Les systèmes de gestion des bases de données ne comportent pas de moyens de protection de la confidentialité, ce qui est très gênant lorsqu'elles sont utilisées à des fins statistiques. De surcroît, les techniques actuelles de prévention de la divulgation ne suffisent pas pour régler tous les problèmes de confidentialité dans les grands entrepôts de données.

25. On ne peut demander aux seuls experts d'un domaine particulier ou aux informaticiens de résoudre ces problèmes. C'est à la direction qu'il incombe de susciter la prise de conscience indispensable face aux problèmes de confidentialité et d'instaurer les préalables qui permettront de trouver des solutions efficaces aux niveaux administratif, méthodologique et technique. Comme certains l'ont fait observer, il est nécessaire d'établir une distinction, dans la politique suivie en matière de confidentialité, entre les données concernant les entreprises et celles qui se rapportent aux personnes physiques. Le problème ne peut être résolu par des moyens techniques. Il faut souvent redéfinir au niveau législatif les règles applicables pour les données relatives aux entreprises.

26. La nécessité d'opter pour un entrepôt de données vaut également pour les organisations internationales. La gestion décentralisée atteint souvent ses limites. Le mode de conception linéaire de la chaîne statistique n'est pas toujours compatible avec la production en commun, et il est souvent déterminé par les impératifs de la production, non par ceux de la diffusion. Eurostat a présenté sa future architecture pour les systèmes d'information, qui est conçue de manière à favoriser l'interopérabilité et la réutilisation des différents systèmes ou applications. Son but est d'instaurer la traçabilité des informations au cours des différentes étapes de leur traitement et de combler l'absence de liens entre différents systèmes.

27. L'utilisation d'entrepôts de données par de nombreux services de statistique et organisations internationales travaillant en réseau pourrait obliger à reconfigurer l'échange de données entre les services nationaux et les organisations internationales. Il faut trouver des solutions efficaces concernant le format dans lequel seront communiquées les données et métadonnées associées. Le rôle des métadonnées dans la diffusion des données est encore plus important aux niveaux international que national. Il pourrait être nécessaire d'élaborer une norme pour l'échange de données et métadonnées statistiques dans un environnement en réseau. Le XML pourrait constituer une solution pour les entrepôts de données et les réseaux. Il pourrait être nécessaire aussi d'élaborer un langage basé sur le XML pour l'échange de données

statistiques. La Réunion de travail sur les métadonnées statistiques tenue en novembre 2000 a lancé une initiative dans ce sens.

28. L'utilisation d'entrepôts de données comme point de départ pour obtenir des données par le biais de l'Internet obligera à mettre au point de nouvelles méthodes et capacités fonctionnelles. D'une part, l'entrepôt de données crée une demande de nouvelles capacités fonctionnelles sur le site Web, et d'autre part le développement rapide de la technologie Internet offre de nouvelles possibilités que pourront utiliser les sites statistiques créés sur l'Internet. Les gestionnaires doivent observer ces évolutions et en tirer des conclusions concernant les tâches à accomplir. Plusieurs participants ont fait observer que les moteurs de recherche sur l'Internet ne sont pas compatibles avec les demandes de renseignements statistiques. Il faut mettre au point des mécanismes particuliers pour rechercher des données statistiques dans les bases de données sur l'Internet.

29. Les principaux destinataires des activités d'un service national de statistique devraient être les utilisateurs, et la création d'entrepôts de données doit se traduire pour eux par des avantages importants. L'application efficace de techniques d'entreposage des données procurera un accès à la fois plus rapide et plus souple aux données. Elle devrait également rendre cet accès beaucoup plus facile. Les méthodes de présentation devraient comporter des visualisations graphiques et géographiques adaptables qui pourraient conduire à une nouvelle qualité de la présentation des statistiques. Une attention de tous les instants est nécessaire pour satisfaire les utilisateurs et connaître leur comportement. Les groupes d'utilisateurs n'ont pas tous les mêmes impératifs et leur comportement évolue avec le temps.

C. Harmonisation des données (d'enquêtes) statistiques avec les données (administratives) issues de registres

30. De nombreux services de statistique utilisent à la fois des données d'enquête et des données administratives. Les enquêtes sont parfois nécessaires pour évaluer ou compléter les sources administratives, qu'il s'agisse des données disponibles, de la qualité des données, de leur cohérence dans le temps, de lacunes à combler ou de différences de concept. En plus des registres administratifs tenus par les autorités publiques, il est possible d'utiliser les données réunies par une organisation pour satisfaire à des besoins administratifs qui lui sont propres, s'agissant par exemple de la tenue des livres comptables ou de l'administration des salaires. Plusieurs pays utilisent un numéro d'identification commun (par exemple le numéro d'identification individuel) pour établir un lien entre différents registres.

31. On constate que divers pays ont tendance à utiliser de plus en plus les registres administratifs à des fins statistiques. Les instruments et pratiques évoluent dans la même direction et les pays qui commencent à mettre en place des systèmes similaires verraient leurs débuts largement facilités si la coopération se développait et que l'on s'efforçait de trouver des solutions qui puissent servir de référence. Il serait utile de parvenir à un accord international sur des définitions et concepts communs afin de pouvoir établir des comparaisons. De nombreux services de statistique envisagent de réunir des données à partir des systèmes d'information des entreprises. Il serait donc utile de s'entendre sur des normes internationales ayant force obligatoire en matière de comptabilité. Par ailleurs, l'harmonisation des règles de traitement des

données dans différents secteurs de la statistique au sein des services de statistique favoriserait de nouveaux progrès dans ce domaine.

32. L'élaboration de statistiques à partir des registres exige de gros investissements au cours de l'étape d'élaboration et peut absorber d'énormes ressources. Cependant, une fois effectué l'investissement initial, le service national de statistique peut réaliser des économies substantielles dans ses opérations au jour le jour. Une telle formule se traduira également par une économie de ressources pour les répondants. Il ne fait aucun doute qu'il faudra à l'avenir faire appel à la fois aux sources administratives et aux sources statistiques pour la production de statistiques. Il n'en reste pas moins de nombreux problèmes pratiques à résoudre pour que cela devienne une réalité.

33. Les sources administratives ont le grand avantage de fournir des informations presque complètes, très homogènes et comportant des données longitudinales. Elles sont généralement très bien acceptées par les unités qui fournissent les informations. Les progrès accomplis dans les domaines de l'informatique et de la technique de mise en relation de fichiers ont grandement facilité une utilisation plus large des registres administratifs à des fins statistiques.

34. Les problèmes rencontrés concernent aussi bien la qualité et la comparabilité des données que les possibilités de réalisation et les aspects juridiques ou encore les ressources. L'obtention de données statistiques à partir de sources administratives exige un travail considérable qui consiste à retraiter les données administratives, à assurer une cohérence globale et à enrichir les données disponibles au moyen d'enquêtes. L'énorme quantité d'informations réunies peut entraîner une lourdeur de gestion et un allongement des délais de traitement, et il est de surcroît compliqué de modifier les questionnaires. La qualité des données provenant des registres dépend dans une large mesure de celle des registres dont elles sont tirées.

35. Les questions de gestion comprennent l'établissement d'une coopération avec les répondants ou les organismes, le recouvrement des coûts et l'adaptation en fonction des changements intervenus dans les sources administratives. Viennent s'y ajouter la gestion des métadonnées et les possibilités de comparaison dans le respect des normes nationales et internationales. Il importe de tenir compte également des ressources humaines ainsi que de la culture et de l'état d'esprit qui y sont associés lorsque l'on s'efforce de modifier l'orientation des statistiques pour faire davantage appel aux données administratives.

36. L'un des principaux problèmes rencontrés lorsque l'on utilise des données administratives tient à leur validité et à leur utilité. Très souvent, ces données ne correspondent pas aux notions statistiques. Dans quelques (rares) cas favorables, les statisticiens peuvent exercer une certaine influence sur le contenu des registres. Un autre problème tient au fait que les changements apportés dans le système juridique et administratif modifie la teneur des données issues des registres administratifs. Les registres répondent à des finalités administratives bien précises et ne peuvent donc satisfaire aux besoins et desiderata dans le domaine statistique. De surcroît, les notions ou catégories de personnes qui n'apparaissent pas dans le registre risquent également de ne pas être pris en compte dans les statistiques. Il importe que certains registres de base répertorient toutes les unités. Lorsque les registres ne peuvent donner un tableau complet de la société, il est nécessaire de réunir des données additionnelles en procédant à des enquêtes. En ce qui concerne l'actualité des données, la situation diffère d'un registre à l'autre.

Certains peuvent rapidement fournir des données très récentes tandis que d'autres, par exemple les registres fiscaux, ne les produiront qu'avec plusieurs mois de retard.

37. La législation peut être source de problèmes lorsqu'il s'agit de préserver le caractère confidentiel de données administratives. La loi interdit parfois l'utilisation de registres en raison du risque de divulgation des données. La Suède, où la divulgation de données à caractère personnel est une infraction aux yeux de la loi, en est un exemple intéressant. Dans ce cas, il est beaucoup plus facile pour le service de statistique de publier des fichiers de microdonnées anonymisées étant donné qu'en cas de violation de la confidentialité, c'est la personne qui a divulgué les données qui sera tenue pour responsable, et non le service qui les a publiées.

38. Les participants ont souligné combien il importait de résoudre les problèmes techniques associés à l'élaboration de statistiques à partir de registres, par exemple la sécurisation des transferts par voie électronique sur l'Internet, la gestion des formats de données et des métadonnées, l'identification correcte des répondants et des unités statistiques ainsi que la gestion de cette identification. Il est nécessaire de disposer d'instruments efficaces pour utiliser, calculer et étayer des données provenant de sources différentes. Il n'existe souvent pas d'instrument dans le commerce et les services mettent au point des solutions qui leur sont propres en fonction de leurs besoins.

39. Parmi les possibilités techniques à exploiter pourraient figurer par exemple l'utilisation du XML, la création de serveurs de schémas XML, la mise à disposition de logiciels pour le commerce interentreprises, par exemple les cybermarchés, ou encore de logiciels d'intégration des applications d'entreprises pour réaliser la transformation et la gestion des données, ainsi que les certificats d'identification sur l'Internet. Certaines initiatives prises dans ce domaine, par exemple l'ebXML, l'UDDI et l'XML/EDI, fournissent un environnement dans lequel il est possible de réaliser des transactions électroniques (registres).

40. Les meilleures pratiques en matière d'élaboration de statistiques à partir de registres comprennent notamment le calcul de données à l'aide de logiciels commerciaux immédiatement utilisables, la création de logiciels de groupe et d'extranets pour faciliter les projets concernant différentes organisations, un répertoire central des métadonnées et un répertoire central pour toutes les données d'entrée éditées.

41. De nombreux participants ont insisté sur la nécessité d'obtenir plus d'informations sur la façon dont les statistiques tirées des registres étaient élaborées dans la pratique dans les pays plus avancés dans ce domaine. Cela faciliterait beaucoup la tâche des pays qui commencent à établir leur système de statistique sur la base des registres. Dans certains pays (les États-Unis et le Danemark par exemple), les meilleures pratiques dans ce domaine sont recensées et expliquées. Les participants se sont déclarés prêts à réaliser un ouvrage méthodologique sur les meilleures pratiques pour l'établissement de statistiques à partir des registres. Les pays nordiques (Danemark, Suède et Finlande, par exemple) et la France pourraient y apporter leur contribution. De l'avis général, cet ouvrage pourrait beaucoup aider les pays, en particulier s'il était associé à un site Internet.

D. Gestion des ressources dans les services de statistique et rôle des sections informatiques

42. Les organismes de statistique doivent réexaminer leurs pratiques en matière de gestion à la lumière des changements qui interviennent dans la technologie de l'information et de l'utilisation croissante de l'Internet. Grâce aux progrès technologiques, les services de statistique peuvent progressivement articuler les opérations de production autour de fonctions telles que la collecte, le traitement et la diffusion des données. Les initiatives en matière d'administration publique en ligne et les innovations dans le domaine de la communication électronique ont elles aussi une incidence sur les tâches des gestionnaires de la TI. Les participants en ont étudié les répercussions sur la gestion des ressources dans les services de statistique et le rôle des sections informatiques. Certaines des solutions proposées pour résoudre les nouveaux problèmes de gestion concernent les systèmes de recouvrement des coûts internes, le travail en équipe et l'informatique répartie.

43. Les services de statistique utilisent différentes stratégies de recouvrement des coûts. L'unité centrale peut par exemple facturer des honoraires pour ses prestations ou appliquer un tarif uniforme par salarié, l'organisme statistique peut faire payer aux clients extérieurs les données diffusées sur l'Internet, ou encore obtenir un financement direct des pouvoirs publics. Le Canada a présenté un exemple de système de recouvrement des coûts internes. Les services de TI sont considérés comme une unité opérationnelle qui recouvre l'ensemble des coûts par la perception de droits pour les services internes et finance également les acquisitions et le remplacement des biens d'équipement.. La TI ne fait pas l'objet d'un financement distinct et le gestionnaire de programme peut donc déterminer la combinaison et la quantité des prestations nécessaires pour atteindre les objectifs de son programme et pour opérer des choix économiques entre fournisseurs internes et fournisseurs externes. Le mécanisme de recouvrement des coûts internes a un avantage, à savoir que ce sont les besoins opérationnels qui encadrent la prestation de services centraux. Les coûts sont directement associés aux projets, ce qui permet de suivre l'évolution des dépenses en ressources humaines et d'opérer des choix stratégiques en cas de compressions budgétaires ou de réaffectation des fonds. Cette formule a modifié radicalement la culture et le comportement de l'organisation. Elle s'est révélée suffisamment souple pour qu'il soit possible de tenir compte des nouveaux besoins au fur et à mesure de l'évolution de la situation. La transparence de l'affectation des fonds engendre la confiance entre les secteurs de service et les programmes.

44. L'organisation du travail en équipes favorise la constitution d'équipes interdisciplinaires et l'utilisation plus efficace du capital humain. La Finlande a présenté une formule de travail en équipes qui remplace l'organisation hiérarchique traditionnelle par une organisation autonome horizontale. Le travail en équipes présente plusieurs avantages : une division plus souple du travail, une meilleure disponibilité des solutions de remplacement, le développement de la coopération et des méthodes de travail interactives, et une plus grande satisfaction procurée par le travail.

45. Le rôle et l'influence du département informatique et de son chef dépendent de la structure organisationnelle du service de statistique. Ils peuvent être très différents selon que le cadre est centralisé ou décentralisé, et il est nécessaire, pour chaque cadre, de mettre au point une méthode qui lui soit propre. On constate à cet égard des divergences dans la démarche adoptée par les

services de statistique. D'une part, le développement du travail en réseau se traduit par le passage à l'informatique répartie. D'autre part, l'entreposage des données et l'intégration des opérations de production impliquent la centralisation des services informatiques.

46. Dans un grand organisme statistique, l'informatique répartie offre de nombreux avantages, tels que l'élargissement de l'accès aux ressources humaines, le développement de la recherche, la stimulation de la créativité et l'accroissement de la flexibilité des ressources financières. Elle présente également des risques; il peut être par exemple difficile de diriger et de contenir l'informatique répartie dans certaines limites s'il n'existe pas de direction centralisée directe, et il devient de plus en plus nécessaire d'assurer une coordination et de normaliser les matériels et logiciels. Le chef de l'unité informatique doit avoir une emprise forte et centralisée sur certaines ressources informatiques qui desservent l'ensemble de l'organisme (l'infrastructure de base et les télécommunications par exemple) et sur certains processus tels que la gestion des données, l'établissement des métadonnées et la recherche-développement. Cela peut avoir une incidence sur la planification stratégique et opérationnelle à travers toute l'organisation. Il est nécessaire de définir des méthodes qui favorisent la convergence des normes, pratiques et contrôles de la qualité.

47. Le chef du service informatique doit s'employer à maîtriser la gestion des ressources informatiques, que le cadre en soit centralisé ou décentralisé. Son rôle devient particulièrement important dans le second cas et il peut être alors amené à instaurer une sorte de "dictature amicale". Pour mener à bien ses tâches de gestion dans un environnement réparti, il doit faire preuve de tout un ensemble de compétences nouvelles, par exemple le sens des relations humaines, la capacité de motiver ses collaborateurs, un état d'esprit privilégiant l'action collective, la vision de l'avenir, l'art de la vente et le don du commandement. Ces compétences vont bien au-delà des qualifications techniques requises traditionnellement pour l'exercice de ses fonctions de gestionnaire. Il importe dans le même temps qu'il se tienne bien informé des dernières innovations techniques et qu'il demeure compétitif.

48. Les défis à relever concernent non seulement la gestion mais également l'acquisition des ressources informatiques. Il est difficile de recruter et de conserver des informaticiens de premier plan car les spécialistes compétents font l'objet d'une surenchère effrénée. Comme les services de statistique ne peuvent pas toujours offrir des rémunérations aussi élevées que dans le secteur privé, il importe de préserver une combinaison appropriée de ressources informatiques en assurant un équilibre entre les prestations externalisées et celles assurées par l'organisation proprement dite. Un autre défi à relever en permanence tient aux ressources financières à acquérir pour continuer à innover et à améliorer les prestations informatiques. La technologie évolue rapidement et le prix des techniques de pointe augmente alors que les budgets demeurent relativement stationnaires. La sécurisation et l'obligation de préserver le caractère confidentiel des données obligent à procéder à des investissements additionnels en pare-feux ou autres mesures de sécurité. Il est bien souvent difficile de justifier le coût des infrastructures. Les crédits sont généralement répartis en fonction des effectifs de personnel sans qu'il soit tenu compte de la complexité relative des domaines informatiques. Il serait utile à cet égard de définir des indicateurs spéciaux et de comparer la situation dans les différents pays.
