



Distr.: LIMITEE  
E/ECA/DISD/CODI.2/14

29 juillet 2001

**NATIONS UNIES  
CONSEIL ECONOMIQUE ET SOCIAL**

**Original: Français**

**COMMISSION ECONOMIQUE POUR L'AFRIQUE**

Deuxième réunion du Comité de l'Information  
pour le développement

Addis-Abeba, Ethiopie  
4-7 septembre 2001

**Rapport sur les Activités Relatives  
au Développement et à la Gestion  
des Bases de Données à la CEA**

## I. Introduction

1. Les premières recommandations faites à la Commission Economique pour l'Afrique (CEA) en vue de la conception, du développement et de l'exploitation d'une base de données statistiques régionales remontent à 1959. Elles suivaient en cela les résolutions adoptées par la première Conférence des Statisticiens Africains pour établir des procédures efficaces de réponse aux différentes demandes en informations statistiques structurées, facilement accessibles et utilisables comme aide à la décision et de facteurs alimentant la recherche aussi bien au niveau de la région qu'en dehors. Plusieurs initiatives sont venues ultérieurement confirmer l'urgence de cette entreprise et enrichir le contenu de la demande en données statistiques sur l'Afrique. La plus importante a été la formulation, par la sixième session de la Conférence conjointe des Planificateurs, Statisticiens et des Spécialistes de la Population et de l'Information, du Plan d'action d'Addis Abeba pour le développement des statistiques en Afrique au cours des années 90.

2. Les travaux de conception et de développement proprement dits n'ont commencé qu'en 1983 et la première structure a été opérationnelle en 1985. L'environnement informatique de l'époque, caractérisé par la prédominance du projet de Système Panafricain de Documentation et d'Information (PADIS), le nombre limité et le manque de qualification dans ce domaine, ont amené à greffer les activités de conception et de développement sur PADIS sous le sigle PADIS-STAT. Ces activités ont, par la suite, été transférées comme un élément de programme prioritaire à la Division de la Statistique, sans jamais avoir bénéficié de l'appui matériel et en hommes nécessaire pour réussir l'entreprise.

3. Avec la mise en place des nouvelles structures de la Commission Economique des Nations Unies pour l'Afrique, le 1er janvier 1997, les activités relevant de la Division de la Statistique ont été intégralement reprises par la nouvelle Division des Services d'Information pour le Développement (DISD). Dans cette nouvelle structure, le développement d'une base de données statistiques régionales a pleinement retrouvé sa dimension multi-sectorielle d'origine en inscrivant dans sa dynamique l'intégration des données en informations aussi bien économiques, que bibliographiques et géographiques.

4. Ce document donne une présentation exhaustive de la structure de la base de données de la première génération développée sur l'ordinateur HP3000, car la migration sur les plates-formes temporaires sous format MS – ACCESS a juste transposé cette structure de base. Cette présentation permet de comprendre les caractéristiques des différents modules existant présentement. Sont également présentées les principales recommandations faites par le consultant chargé de l'étude de faisabilité pour le développement du Système d'Information Statistique de la

CEA (SIS-CEA), ainsi que les activités similaires entreprises dans d'autres Divisions opérationnelles de la CEA, y compris les Centres sous-régionaux de développement (CSRD), et avec d'autres partenaires extérieurs.

## **II. La base de données statistiques de première génération**

5. Cette base de données a été conçue et développée dans le cadre du projet PADIS. L'environnement de calcul et les ressources humaines et financières disponibles à l'époque étaient relativement limités, comparé à l'ampleur de la tâche à exécuter. La base de données a été installée sur l'ordinateur HP3000/58, disponible à l'époque, avec une mémoire principale de 5 mégabytes, 4 lecteurs de disques ayant une capacité de stockage de 404 mégabytes chacun, 2 dérouleurs de bandes magnétiques d'une densité d'enregistrement de 1600 et de 1600/6250 bpi, des imprimantes et des terminaux à écran de visualisation. Elle était accessible à partir de tous les terminaux connectés au système dans l'enceinte de la CEA et à partir du siège de l'OUA par l'intermédiaire d'une ligne téléphonique spéciale.

6. L'ordinateur HP3000 était géré par le système d'exploitation MPE4, alors que les logiciels de gestion de fichiers comprenaient : IMAGE 3000, QUERY 3000, KSAM. Les programmes utilitaires comprenaient : INSIGHT, VIEW, TDP (traitement de texte), SORT, MERGE et FCOPY. Les langages de programmation BASIC, COBOL, FORTRAN, RPG et SPL ont servi à la programmation des différentes composantes.

7. Les travaux de conception, développement et maintenance de la base de données ont été conduits par une équipe restreinte constituée par :

- 1 informaticien à plein temps ; et,
- 2 statisticiens à temps partiel pour le développement de la base de données, les autres statisticiens de la Division intervenant pour la gestion de leurs données.

8. La première base de données statistiques a été conçue en trois niveaux intégrés qui devaient constituer le noyau d'un système global d'information statistique pour la saisie, le stockage, la gestion, le traitement, et la diffusion des données statistiques. Le premier niveau donne une image succincte, par monographies de 2 à trois pages, de la situation économique et sociale de chacun des Etats membres de la Commission. Le second contient un ensemble de données statistiques détaillées au plus bas niveau de désagrégation. Il devait servir de principale source de données de base pour les deux autres niveaux. Son développement devrait être la première étape du processus d'établissement de tout le système. Cependant, cela n'a pas été possible en raison de l'insuffisance des moyens de traitement électronique des données et du

manque de personnel qualifié. Sous sa forme installée sur l'ordinateur HP3000, ce niveau ne contenait que les données du commerce international et celles de la comptabilité nationale stockées sur bandes magnétiques pour un traitement ultérieur groupé. Le troisième niveau est formé de séries chronologiques portant sur 13 secteurs différents, à savoir : l'agriculture, les transports et communications, l'éducation, les finances, la santé, l'industrie, l'emploi, la population, le commerce international, les statistiques des faits d'état civil, les prix, l'environnement et les comptes nationaux. Ces séries sont celles qui sont fréquemment demandées par les utilisateurs. Cependant les contenus sectoriels ont connu des niveaux de développement inégaux. Cette composante du système d'information a été également développée en priorité en raison de la forte demande interne en séries chronologiques et la nécessité d'accélérer le processus d'édition des publications statistiques de la CEA.<sup>1</sup> La capacité initiale avait été fixée à 150000 entrées, pour 120000 séries. Les langues d'utilisation prévues étaient le français et l'anglais. Cependant seule cette dernière a été opérationnelle.

9. La structure logique est présentée dans le tableau 1 donné en annexe. La codification a été faite, dans la mesure du possible, selon des nomenclatures internationales : la Classification Type pour le Commerce International (CTCI) pour les séries du commerce international ; la liste partielle des produits et matières dérivés de la Classification Internationale Type par Industrie de toutes les branches d'activité économique (CITI) pour les statistiques industrielles ; le Répertoire des Statistiques Internationales ; et le « Standard Country Code » de la Division de Statistique des Nations Unies. Des codes ont été introduits pour couvrir les regroupements sous-régionaux et régionaux. La structure physique des données suit les règles applicables aux bases de données gérées par IMAGE, à savoir une organisation des fichiers en deux ensembles : un ensemble principal dit « maître », pouvant être manuel ou automatique, et ensemble dit « détails », qui est une collection des enregistrements relatifs aux données. Le premier permet l'accès au second par l'intermédiaire des identifiants.

10. La collecte des données de base sur les pays membres par la CEA n'est ni systématique, ni régulière. Elle se fait présentement sur une base ad-hoc et en fonction des besoins des publications annuelles de la Commission, à savoir l'Annuaire Statistique, le Répertoire des Statisticiens Africains ou des Centres Africains de Traitement des Données, le Rapport Economique et l'Etude des conditions économiques et sociales en Afrique. La collecte se fait par l'envoi de questionnaires, de missions auprès des pays membres (pratique très limitée en raison de la rareté des ressources financières) et par l'exploitation des publications d'autres agences et organisations spécialisées, telles que la FAO, l'ONUDI, l'OCDE, la CNUCED et la Division de Statistique des Nations Unies (UNSD). Le système de validation est décentralisé au niveau des postes de travail des aides statisticiens attachés aux différents domaines de la base de données.

11. Le troisième niveau, qui a été le plus opérationnel, a été conçu et développé avec le système de gestion de base de données IMAGE3000 qui est un produit Hewlett-Packard. IMAGE3000 est utilisé pour définir et créer la base de données et y accéder. Ce système de gestion de base de données général n'est pas adapté aux besoins particuliers des statisticiens. Des programmes spéciaux ont donc dû être développés pour adapter ses caractéristiques à ces demandes spécifiques, concernant notamment : le traitement des données statistiques, la création de tableaux, les calculs statistiques de base, la maintenance de la base de données et la mise à jour régulière des données dérivées.

12. Cette base de données a servi à mettre en place un embryon de système rationnel pour l'édition des publications statistiques : Annuaire statistique pour l'Afrique, Indicateurs socio-économiques, Statistiques du Commerce extérieur, Annexes statistiques de l'Etude des conditions économiques et sociales en Afrique et divers autres tableaux produits et envoyés aux pays et aux organisations internationales sous des formats variés. Cependant, son exploitation a été limitée par les contraintes suivantes : (a) le dépassement des espaces disque disponibles en raison de la grande taille des séries ; le manque de portabilité qui a limité son implantation dans les pays membres ; (c) la lenteur de l'édition des tableaux statistiques ; (d) les procédures de mise à jour contraignantes en raison de la lourdeur des modes d'accès ; (e) la marginalisation des pays membres dans le processus de conception, de développement et même de mise à jour ; (f) les risques fréquents d'erreurs dans les agrégations qui ne sont malheureusement pas calculées au moment de l'édition des tableaux, mais au préalable et entrées dans la base ; (g) les possibilités de consultation en mode interactif limitées en ce sens qu'elles ne permettent que des consultations série par série et non par lots ; (h) l'inexistence d'une procédure de contrôle des données elles-mêmes ; et, (i) l'absence d'un système de méta-données intégré à la base. Il faut également citer les possibilités limitées de fonctions de calculs et d'affichage de graphiques. Ces contraintes et la persistance du manque de personnel qualifié ont fait que les niveaux I et II n'ont jamais fait l'objet de développements ultérieurs.

### **III. De l'ordinateur HP3000 aux ordinateurs individuels : la migration**

13. Au milieu des années 90, l'ordinateur HP-3000/58 a atteint la limite de sa durée de vie et de sa capacité de stockage. Les interruptions dans le fonctionnement sont devenues fréquentes. Le secrétariat a dû entamer sa mise hors circuit. Dès lors, il était devenu impérieux de transférer les données résidentes sur une plate-forme, même transitoire, en attendant que les infrastructures physiques nécessaires au développement d'un réseau opérationnel soient mises en place et que les compétences pour la conception et le développement de la base de données, des interfaces et applications requises soient disponibles.

14. Avec la généralisation des ordinateurs individuels (PC) au niveau du Secrétariat, une bonne partie des données stockées sur l'ordinateur HP 3000 a été transférée sur les postes de travail des assistants de recherche en attendant la migration de l'ensemble de la base de données statistique. Cependant, ces transferts ont été opérés sous différents formats, QUATTRO, PARADOX et EXCEL, et de manière dispersée, avec un manque de coordination et de cohérence dans les normes et standards de classification et dans les procédures de validation. Le développement du réseau local (LAN) et la mise à disposition d'un serveur pour les activités statistiques ont permis d'entreprendre la migration de la base de données de manière intégrée sur le réseau.

15. Pour cette phase, le modèle de données et le système de gestion de base de données relationnelles ont été choisis comme standard et plate-forme de migration et comme logiciel de gestion de la base de données. Ce choix semble justifié, car il est le plus adapté à un environnement de calcul bâti sur des ordinateurs individuels (PC) tournant en réseau et utilisant MS-Windows comme système de gestion, comme c'est le cas à la CEA. En outre, le personnel chargé de la maintenance a acquis de solides connaissances sur ce logiciel, ce qui permet une économie de temps et de ressources pendant cette phase transitoire.

16. Cependant, le volume important et la taille des séries rendent difficiles les manipulations de la base entière en utilisant les ressources et les outils de MS-ACCESS. La base initiale a ainsi été décomposée en 12 bases de données sectorielles de taille plus modeste et utilisable sur des ordinateurs individuels. Il s'agit des composantes relatives aux secteurs suivants (une base de données pour chacun d'entre eux) : agriculture, transport et communications, population et statistiques sociales, finances, industrie, commerce, prix, comptes nationaux, énergie, environnement et tourisme. Les modules énergie, environnement et tourisme transférés à partir de l'ordinateur HP3000 sont vides de tout enregistrement, comme c'était le cas pour les fichiers d'origine. Cette situation persistera si le taux de réponses aux questionnaires envoyés aux Etats membres sur ces deux secteurs n'augmente pas. Les seules ressources disponibles en ce moment sont : les sites Web nationaux existants, l'Annuaire international des statistiques du tourisme, les sites de la Banque Africaine de Développement et celui d'AFRISTAT. Les classifications utilisées pour la codification de l'ensemble des données sont les mêmes que celles de l'ancienne base de données.

17. La base de données centrale ainsi que ses composantes sectorielles, en format MS-ACCESS, ont été ainsi conçues autour de six ensembles de tables à la structure semblable à celle de la base de données sur l'ordinateur HP3000. Il y a une table détaillée de données statistiques (« detail statistical data table »), avec 6289 enregistrements et cinq tables principales de données statistiques (« master data set tables ») de 122007 enregistrements. Ce dernier ensemble comprend

le code de classification des séries statistiques, les modes d'observation des séries, les codes des pays déclarants, le nom des pays et la monnaie y ayant cours, les valeurs observées des séries et la table des unités de mesure. Les données de cette partie sont disponibles aussi bien sous forme textuelle que sous format MS-ACCESS. En outre, la conception et la structure de toutes les bases de données sectorielles et de leurs tables principales sont identiques. Ceci facilite aussi bien le traitement des modules sectoriels en postes indépendants que l'intégration des composantes le moment venu.

18. Des menus très conviviaux ont été créés dans les utilitaires de MS-ACCESS, formulaires, requêtes, états et macros, et des procédures développées pour une exploitation facile des bases de données sectorielles. Il en est de même pour la génération des tableaux-types pour l'Annuaire Statistique, le Recueil des Statistiques du Commerce extérieur, et pour les procédures d'agrégation.

19. Les problèmes et contraintes suivants sont apparus au niveau général pendant la migration. En effet, dans un premier rapport intérimaire sur le développement de la base de données statistiques régionales de la CEA, le nombre d'enregistrements était estimé à 160000, avec des séries couvrant la période 1965-1990. Ce nombre n'était que de 120000 au moment de la migration et la période couverte se limitait à 1970-1995. La CEA a donc perdu un nombre important de données. Les données sur le commerce extérieur qui avaient été temporairement stockées sur plus de 370 bandes magnétiques ainsi que la documentation sur la base de données elle-même ont disparu. Cependant, les données disponibles peuvent être utilisées avec différents types de systèmes de gestion des bases de données (SGBD) et applications pour développer le noyau du système d'information de la CEA.

20. Au niveau des composantes sectorielles, une certaine mise à jour des normes et standards s'impose. Il s'agit, entre autre : de la conversion des séries de comptes nationaux du Système de Comptabilité Nationale de 1968 (SCN 68) au SCN 93 ; d'une nouvelle écriture des programmes de traitement, notamment pour les comptes nationaux et les indices de prix, et/ou l'acquisition de logiciels appropriés, comme ERETES.

21. Les données accessibles les plus récentes pour les transports et communications datent de 1997 - sauf pour de très rares pays - et d'importantes différences subsistent entre les différentes sources, nationales et internationales. Les tentatives faites pour améliorer la couverture ont été vaines, car le taux de réponse aux questionnaires envoyés aux Etats membres n'a été que de 23 %. Le secrétariat dépend donc pour ce secteur des sources internationales telles que : la Division de Statistique des Nations Unies (publications de statistiques sur les transports routiers, les chemins de

fer et les communications), l'UNESCO pour les infrastructures de communications, la Banque mondiale à travers sa base de données sur les chemins de fer. La couverture des statistiques maritimes et des transports aériens reste très pauvre. Il est également important d'élargir le contenu de la base de données pour tenir compte des nouvelles innovations technologiques comme le téléphone portable, l'accès à Internet, etc.

22. L'essentiel des statistiques de prix disponibles repose sur l'indice des prix à la consommation, et souvent seulement la consommation en milieu urbain. Les données les plus récentes dataient de 1995/96 mais des séries correspondantes à jour ont été extraites pour la plupart de l'Annuaire des Statistiques du Travail publié par le BIT, des sites web nationaux et du site web d'AFRISTAT, où des séries longues de 30 ans sont publiées pour l'ensemble des Etats membres de l'Observatoire. Les statistiques de prix sont celles qui sont le plus régulièrement publiées dans la plupart des pays, très souvent sur une base mensuelle. Ces ressources en données sont assez largement disponibles sur Internet et la mise à jour de cette composante de la base de données ne pose pas de problèmes particuliers. Le secrétariat a entrepris une collecte exhaustive des données nationales dans ce secteur, qui est un des exemples de réussite pour la publication régulière de statistiques nationales fiables et à temps. Enfin, la maintenance des modules relatifs aux secteurs énergie, industrie et commerce international a pris du retard, en raison de l'indisponibilité des données récentes et des difficultés d'accès aux sources et ressources extérieures de données, notamment le Centre International de Calcul de Genève.

23. En conclusion, la multiplicité des postes de validation et stockage des données provenant de l'ordinateur HP3000 et de maintenance des composantes sectorielles de la base de données actuelle, le manque de documentation (problème des métadonnées) et les difficultés d'accès aux sources extérieures ainsi que les problèmes d'accessibilité aux données disponibles, surtout le manque d'accès en ligne pour les utilisateurs, sont les problèmes majeurs qui retardent le développement d'une base de données statistiques fiable et opérationnelle. Un système de méta-données et d'informations sur la fiabilité et les méthodologies de compilation doit donc être intégré à la base de données ainsi qu'à ses différentes composantes. La promotion d'une utilisation intensive des ressources disponibles sur Internet et des échanges électroniques de données entre le secrétariat, les Etats membres et les institutions internationales, surtout celles qui ont développé des systèmes de méta-données pour accompagner leur production statistique doit être faite. Les activités actuelles de l'équipe chargée du développement, de la maintenance et de la gestion des bases de données sont orientées dans ce sens.



#### **IV. Les bases de micro-données**

24. Le secrétariat a également développé des bases de données thématiques à partir de fichiers d'Enquêtes Démographiques et de Santé (EDS/DHS-Africa). Ces fichiers constituent des instruments privilégiés pour permettre la comparaison et l'harmonisation de statistiques socio-démographiques et de santé ainsi que pour développer une approche non monétaire de la pauvreté. Ils contiennent des indications sur les pratiques et l'accès à la santé, les comportements vis à vis des maladies sexuellement transmissibles dont le VIH/SIDA, la fertilité et ses déterminants et de nombreuses variables socio-démographiques. Des bases de données ont été construites pour 29 pays africains pour lesquels un total de 48 enquêtes ont été menées et 82 indicateurs définis, aussi bien en milieu urbain que rural. Elles sont diffusées sur des CD-ROM.

25. Le contenu des bases de données a été élargi à 136 indicateurs et d'autres thèmes de structuration ont été ajoutés à la structure initiale. Les métadonnées présentées ont également été complétées. Le travail actuellement en cours porte sur des tests de comparabilité et d'harmonisation des indicateurs basés sur 15 à 20 fichiers d'enquêtes de types différents. Une présentation de ces bases de données aura lieu au cours de cette session.

#### **V. La tendance actuelle : une base de données sur client-serveur et bâtie sur les besoins des utilisateurs**

26. Les réformes intervenues au sein du secrétariat ont abouti à la réorientation des activités, notamment statistiques, vers les besoins réels des utilisateurs. Ce réajustement en termes de développement des bases de données a été facilité par les facteurs suivants, entre autres : (a) la promotion de l'utilisation intensive des nouveaux concepts et outils de développement des bases de données tels les nouveaux systèmes de gestion de base de données relationnelles et la programmation orientée objet ; (b) l'expansion des réseaux et le partage des ressources en données et méthodologies statistiques ; et, (c) la tendance à la généralisation de l'utilisation des outils et ressources d'Internet dans les activités statistiques. C'est ainsi que le secrétariat a lancé la conception et le développement de son système d'information statistique, dont sa base de données régionale multi-sectorielle.

27. Les premières phases ont été achevées, avec :

(a) l'inventaire de l'existant, (système d'information, base de données et environnement informatique), au niveau du secrétariat et des CSRD qui a été conduit à travers des interviews auprès des différents acteurs ;

(b) la conception des projets de modèles conceptuels de données (MCD) et de modèles conceptuels de traitements à partir du sous ensemble représentatif identifié à partir de l'évaluation de l'existant ; et,

(c) l'architecture fonctionnelle du nouveau système et la planification rigoureuse de sa mise en oeuvre et du scénario retenu pour la conception et le développement du système d'information de la CEA, y compris l'environnement technique, les plates-formes alternatives, les logiciels, le SGDR, les serveurs, les modes de diffusion et les coûts (humains, matériels, financiers).

28. On peut déjà signaler les insuffisances du système actuel relevées lors de l'évaluation de l'existant en termes de données et traitements :

(a) la multiplicité des sources de données qui expose les utilisateurs à des choix difficiles et entraîne la publication de statistiques contradictoires et inconsistantes au niveau du même secrétariat ;

(b) le cloisonnement des divisions, i.e, l'absence d'échanges internes d'informations qui détourne les utilisateurs, même ceux de la CEA, vers les sources extérieures et crée un manque de coordination dans les mises à jour ;

(c) l'accessibilité limitée aux données sensibles, notamment certaines données d'enquêtes auprès des ménages qui sont essentielles pour l'analyse des problèmes de pauvreté ;

(d) l'insuffisance, voire l'absence, des flux de données entre le siège et les CSRD ;

(e) la duplication des traitements et le manque de partage et de rationalisation des ressources ; et,

(f) les mises à jour non coordonnées.

29. Des recommandations ont également été faites sur les aspects gestion, organisation et orientations techniques. Le rapport final du consultant, dont les termes de référence sont donnés dans l'annexe 2, sera disponible sous peu.

## **VI. Conclusions et recommandations**

30. La préservation des données existantes sur une plate-forme transitoire, l'inventaire de l'existant en termes de bases de données et systèmes d'information, l'identification des besoins

des utilisateurs et les propositions techniques d'architecture fonctionnelle constituent une base solide pour les développements futurs du SIS-CEA, surtout après la levée des contraintes de ressources et en personnel qualifié. La planification de la mise en oeuvre doit être flexible pour permettre des choix techniques en adéquation avec les besoins des utilisateurs et de l'environnement existant déjà. Le succès de l'entreprise passe par la mise sur pied d'un comité multi-disciplinaire de suivi, formé de fonctionnaires des différentes divisions opérationnelles.

**Annexe 1**  
**Structure logique des données**  
**de la base de données statistiques de première génération**

<b>Variables</b>	<b>Description</b>	<b>Longeur en bytes</b>	<b>Type</b>
<b>SERCODE</b>	Identificateur de série	6	Alphanumérique
<b>REPCODE</b>	Code du pays déclarant	2	Alphanumérique
<b>UNIT</b>	Unité d'observation	2	Numérique
<b>MODE</b>	Mode d'observation (...)	2	Alphanumérique
<b>PARCODE</b>	Code pays partenaire	4	Numérique
<b>BASEYR</b>	Année de base de la série, le cas échéant	2	Numérique
<b>PRDTYPE</b>	Type de période (...)	2	Alphabétique
<b>STARTYR</b>	Année début série	2	Numérique
<b>YRT</b>	Valeur d'observation pour l'année T	9	Numérique signé
<b>FNT</b>	Type et source de la donnée pour l'année T	2	Alphanumérique
<b>UPDATE</b>	Date de mise à jour	6	Numérique
<b>CRID</b>	Identificateur du créateur	2	Alphabétique
<b>DATATYPE</b>	Indicateur de confidentialité des données	2	Alphabétique
<b>EXPDATE</b>	Date d'expiration pour les données confidentielles	6	Numérique

## **Annexe 2**

### **Conception et développement du système d'information statistique de la CEA**

#### **Termes de référence**

##### **Activités à entreprendre**

1. Organiser des entretiens approfondis avec les fonctionnaires et les organes de direction de la Commission Economique pour l'Afrique (CEA) afin d'identifier et articuler leurs besoins fonctionnels en matière d'information pour le développement et de traitement des données ;
2. Identifier les principaux flux d'information et relations fonctionnelles ainsi que les utilisations et utilisateurs actuels et potentiels de l'information statistique de la CEA ;
3. Sur la base des paragraphes 1 et 2, définir un ensemble complet de composantes nécessaires à la mise en place d'un système informatisé d'information statistique incluant, entre autres, la saisie, le stockage et la gestion des données, ainsi que la description des modèles de données ;
4. Compte tenu de l'environnement de calcul en place à la CEA (logiciels et machines), du point de la situation en matière de technologies de l'information et de bases de données, ainsi que des considérations de coûts, proposer un choix de matériel et de système de gestion de base de données (SGBD) et d'interface graphique les mieux adaptés à la CEA et dont l'entretien se réduit au minimum ;
5. Consolider les paragraphes 1 à 4 en un projet technique cohérent d'architecture fonctionnelle du système d'information ;
6. Etablir le programme d'exécution et le devis de ce projet ;
7. Combiner les points 1 à 6 en un rapport final. Ce rapport doit être rédigé en français puis traduit en anglais et présenté sur support informatique en format MS-Word. Il doit comprendre, entre autres :
  - (a) un résumé analytique ;
  - (b) une analyse des besoins des utilisateurs, tenant compte, notamment, des éléments suivants :
    - (i) inventaire de la distribution et du contenu des bases de données et systèmes d'information actuelles et à venir au niveau de la CEA et des Centres sous-régionaux de développement ;
    - (ii) répertoire, consultable en ligne, indiquant le type et la localisation des données disponibles, ainsi que les modalités d'accès et de manipulation ;
    - (iii) interfaces utilisateurs ;

- (iv) saisie des données, y compris les procédures de saisie et de vérification, et la vitesse de traitement de la saisie à la production des rapports ;
  - (v) clauses de confidentialité et de sécurité ;
  - (vi) protocoles de stockage et de gestion ;
  - (vii) possibilités de recherche en ligne ;
  - (viii) production des rapports ;
  - (ix) fonctions d'analyse statistique et graphique ;
  - (x) fonctions de projection, de modélisation et de mappage ;
  - (xi) fonctions d'importation et d'exportation des données ;
- (c) Un projet technique d'architecture fonctionnelle du système, incluant, entre autres :
- (i) le degré de centralisation des données ;
  - (ii) le degré de flexibilité et d'adaptabilité des capacités de stockage ;
  - (iii) les interface utilisateurs ;
  - (iv) les différentes procédures de saisie ;
  - (v) les différentes procédures de recherche, de traitement, de production des rapports, d'analyse et de présentation des données ;
  - (vi) les possibilités d'échanges informatisés de données ;
  - (vii) les besoins en ressources humaines ;
  - (viii) les possibilités d'aide en ligne.
- (d) Un diagramme des composantes du système assorti de propositions de logiciels ;
- (e) Une analyse comparative du système recommandé et des systèmes actuels de la CEA et de ceux qui sont disponibles sur le marché ;
- (f) Une estimation des coûts et du calendrier de mise en place de ce système.
8. Présentation du rapport à un Comité d'évaluation formé des fonctionnaires de DISD et d'autres provenant des divisions qui ont été enquêtées au cours de la phase I.
9. Evaluation du rapport et ajustements éventuels en fonction des commentaires reçus.

**Qualification**

Diplôme universitaire en informatique du niveau maîtrise (ou plus), avec des connaissances avancées en statistiques.

**Expérience**

Cinq à huit années d'expérience en programmation informatique et dans la conception et le développement des systèmes d'information au niveau national et international (de préférence dans des structures d'enseignement et de recherche), ainsi que dans la conception des logiciels de gestion des bases de données. Connaissance approfondie des réseaux et des systèmes d'exploitation UNIX et NOVELL NETWARE. Aptitude à l'enseignement.

**Aptitudes linguistiques**

Français et/ou anglais parlés et écrits couramment.

**Durée :** 2 à 3 mois

<b><u>Coût :</u></b>	1.	<u>Honoraires (2 à 3 mois)</u>	US\$ 18,000 (US\$ dix huit mille)
	2.	<u>Voyage et per diem</u>	US\$ 6,700 (US\$ six mille sept cents)
		<u>Total</u>	US\$ 24,700 (vingt quatre mille sept cents)

**Mode de paiement :** Mensuel après évaluation de l'état d'exécution de chaque étape/phase

**Date début contrat :** 09 Avril 2001.

**Notes**

1. Les principaux fichiers de la base sont : (a) le fichier schéma (SHEMA) ; (b) le fichier des séries (SERIES) ; (c) le fichier des modes d'observation (MODES) ; (d) le fichier des pays déclarants (REPORTER) ; (e) le fichier des pays partenaires pour le commerce international (PARTNERS) ; (f) le fichier des unités d'observation (UNITS) ; et , (g) le fichier des mesures des séries (SDATA). Trois autres fichiers complètent le système pour : (h) informer sur l'identification des créateurs de séries (USER) ; (i) vérifier les mots de passe et les modes d'accès (USERINDEX) ; et, (j) contrôler le processus des transactions en mode conversationnel (PASSWORD-INDEX). Un fichier séquentiel indexé (KSAM) a été incorporé au système pour accélérer l'impression des tableaux en traitement par lots.