Distr.
GÉNÉRALE

CES/AC.36/1998/14
EUR/ICP/INFO 020603/14
24 juillet 1998

FRANÇAIS
Original : ANGLAIS

COMMISSION DE STATISTIQUE et
COMMISSION ECONOMIQUE POUR L'EUROPE

ORGANISATION MONDIALE DE LA SANTÉ
BUREAU RÉGIONAL DE L'EUROPE

CONFÉRENCE DES STATISTICIENS EUROPÉENS

Réunion commune CEE-OMS sur les statistiques de la santé
(Rome, Italie, 14-16 octobre 1998)

SESSION II : Rassemblement d'informations relatives à la santé - apport de l'informatique

LE SYSTÈME ALLEMAND D'INFORMATION SANITAIRE

Document d'appui présenté par l'Office fédéral des statistiques de l'Allemagne 1/

Résumé

1. Le *Système allemand d'information sanitaire* a été créé dans le cadre du projet "Création d'un Système fédéral de surveillance de la santé". Le contenu et la structuration de ce système reflètent les principes sur lesquels le projet de recherche est fondé.

2. L'objectif du projet était de présenter des données détaillées sur la situation sanitaire de la population allemande et les facteurs déterminants, dont les services de santé. Toutes les données sont groupées selon sept thèmes majeurs, à savoir :

- Structuration du système des services de santé,
- Situation sanitaire
- Les comportements et autres facteurs de risques pour la santé
- Les maladies
- Ressources du secteur de la santé
- Production et consommation des services de santé
- Dépenses de santé et coûts et financement des services de santé.

1/ Rédigé par Gunter Brückner, Service de surveillance de la santé.

3. En tant que produit dérivé de ces activités, le Système d'information sanitaire avait pour principal motivation la mise en place d'une infrastructure commune de données cohérentes et faciles à utiliser pour les collaborateurs de l'élaboration du premier "Rapport sur la santé en l'Allemagne".

4. Assez rapidement il est apparu évident que ce trésor de données, à ce jour encore jamais rassemblées de cette manière centralisée, ne pouvait rester inexploité une fois le rapport terminé. Devant le succès grandissant d'Internet, nous avons eu l'intuition du parti que nous pourrions tirer de ce réseau pour rendre ces informations accessibles à tous. Un système d'information ayant pour support Internet présente un double avantage, à savoir, un service commode aux utilisateurs, et une rentabilisation de ressources humaines limitées.

5. Il fallait résoudre des problèmes majeurs avant de rendre ce système accessible au public. La présente communication expose ces problèmes et les solutions trouvées. Il traite de la structuration, de la modélisation des données, des matériels et logiciels utilisés, de la conception ergonomique appliquée au système pour le rendre convivial et des problèmes de sécurité associés à un tel système dans un environnement Internet-Intranet. Le *traitement analytique en ligne* (TAL) est la pierre angulaire de la technologie d'information utilisée pour ce projet.

6. Nous pensons que les principes de base de ce système pourraient bien servir de tremplin à la communication d'informations qui soient comparables au niveau international dans les domaines touchant à la santé.

I. Introduction

1. En 1992, le Ministère fédéral de la recherche-développement et le Ministère fédéral de la santé ont lancé, conjointement, un projet de recherche sur le thème "création d'un système fédéral de surveillance de la santé". Le but était de rendre compte, pour la première fois et de manière détaillée, de l'état de santé de la population allemande et des facteurs déterminants, dont les services de santé. Il a été explicitement prévu d'intégrer les paramètres épidémiologiques et économiques. Parallèlement, le projet devait préparer la voie aux futures activités régulières d'information dans le domaine de la santé.

2. Lorsque ce projet a été conçu , il était notoire que la situation laissait à désirer : les données étaient dispersées entre de multiples "propriétaires", leurs définitions étaient incohérentes, elles n'étaient pas normalisées et bien souvent il était difficile de les retrouver. Les indicateurs étaient calculés et recalculés par divers organismes qui n'obtenaient pas toujours les mêmes résultats. Le travail des chercheurs était fastidieux du fait qu'il leur fallait commencer par établir qui était en possession de telles ou telles données et à quelles conditions elles leur seraient communiquées.

3. En quatre ans, nous avons réussi à nous assurer la coopération de tous les fournisseurs. Ils ont pu se rendre compte que cette collaboration leur était profitable en facilitant leur tâche. Le projet suscitait beaucoup d'espoirs, aussi fallait-il tout entreprendre pour répondre à cette attente.

II. Structuration

4. Les concepteurs du projet de recherche "Création d'un système fédéral de surveillance de la santé" visaient à atteindre les objectifs suivants :

- Établir le premier "Rapport sur la santé en Allemagne" traitant en une centaine de chapitres d'une variété de thèmes ayant trait à la santé. Tous ces thèmes devaient être traités dans le détail et de manière cohérente et couvrir les aspects suivants : "État de santé", "Facteurs de risque pour la santé", "Maladies", "Services de santé", "Ressources du secteur de la santé" et "Dépenses, coûts et financement". Chaque thème devait être traité à partir de bases empiriques et être présenté sous une forme compréhensible pour le plus grand nombre.

- Établir au moins deux rapports thématiques sur des questions de santé prêtant à controverse ou hautement sensibles d'un point de vue politique. Les thèmes "Allergies" et "Soins de santé ambulatoires et internes interactifs" ont été choisis.

- Créer un Centre de documentation et d'information sur les données sanitaires (Information and documentation center "Health Data" - IDC) qui servirait de base de données pour toutes les activités de communication.

- Exécuter des projets de recherche méthodologique ad hoc aux fins d'élargir ou d'améliorer la base de données empiriques ou la présentation des données. Les thèmes "Un nouveau concept pour le calcul des dépenses de santé", "Coûts de la maladie", "Recherche d'informations sur les soins de santé, par type de maladie à partir des données sur les dépenses de santé de la population" ont été choisis.

- La méthodologie suivie devait permettre un prolongement des activités sur une base régulière. L'intégration facile de nouveaux thèmes et la possibilité de modifier l'ordre des priorités sans difficulté devant être des préoccupations majeures.

5. Ce souci de pérennité, l'ampleur des éléments couverts et le caractère plus ou moins détaillé de ces éléments exigeaient une modélisation des données du système d'information extrêmement souple.

6. Les nouvelles techniques d'information devenant de plus en plus largement accessibles, grâce surtout à Internet, les utilisateurs se sont montrés de plus en plus exigeants. Les moteurs de recherche Internet sont devenus des repères de référence majeurs et les utilisateurs se sont de plus en plus habitués à la recherche par mot clé, à l'extraction intégrale de

texte, à la recherche en chaîne par élimination et aux systèmes multilingues. Le système d'information sanitaire de l'Allemagne se devait d'être à la hauteur de ces réalisations.

7. Nous étions par ailleurs persuadés qu'un tel système d'information réglerait une fois pour toutes les problèmes associés au calcul de la valeur des indicateurs. Tout spécialiste de l'information s'est trouvé confronté aux difficultés évoquées ci-après.

8. Quel que puisse être le nombre d'indicateurs calculés et stockés sous forme de tables, le premier client venu ne manquera pas de demander un arrangement ou une ventilation qui n'a pas été prévu(e) alors même que sa demande est légitime et peut être satisfaite sur la base des données existantes.

- La révision d'un nombre dans une source de données peut influencer une énorme quantité de valeurs. Il en découle, soit que tous les indicateurs doivent être recalculés soit qu'on accepte le fait que leurs valeurs ne concordent pas.

- L'introduction d'un nouvel indicateur est la cause de problèmes insurmontables en ce qui concerne les données antémémorisées. Dans la plupart des cas, les valeurs de ces indicateurs ne peuvent être établies pour des périodes antérieures.

- Les doublets sont chose commune. Personne ne peut estimer le nombre de personnes qui, par exemple, "programment" les tables de mortalité du progiciel de numérisation, ni prévoir si ces tables produiront des résultats comparables.

III. Élaboration du système d'information sanitaire

III.1 Principes de base

9. Cette analyse des problèmes et de la structuration a conduit aux principes directeurs ci-après :

- Tout système d'information doit tenir compte du fait que les informations se présentent sous forme textuelle et numérique. Les textes et les nombres communiquent des informations, chacun d'une manière spécifique. Les nombres peuvent se présenter sous forme de tables ou de diagrammes. Les textes peuvent contenir des tableaux ou des graphiques à titre d'illustrations. L'information peut avoir différents ports, de nos jours les plus importants sont les imprimantes et les dispositifs de lecture électronique. L'information est fournie en réponse à une requête. Il ne peut donc y avoir de réponse s'il n'y a pas de requête, d'où il découle que tout système d'information doit pouvoir lire ou isoler la requête de l'utilisateur faute de quoi il ne lui sera pas possible de fournir la réponse appropriée, même s'il contient toutes les données nécessaires.

- *Le texte* est l'instrument de communication le plus important entre l'utilisateur et le système, aussi doit-il décrire clairement l'objet recherché. Cela est particulièrement important en ce qui concerne les nombres : peut-on attendre d'un utilisateur qu'il recherche, par exemple "81.785.625" pour obtenir la réponse "Population allemande, 1995".

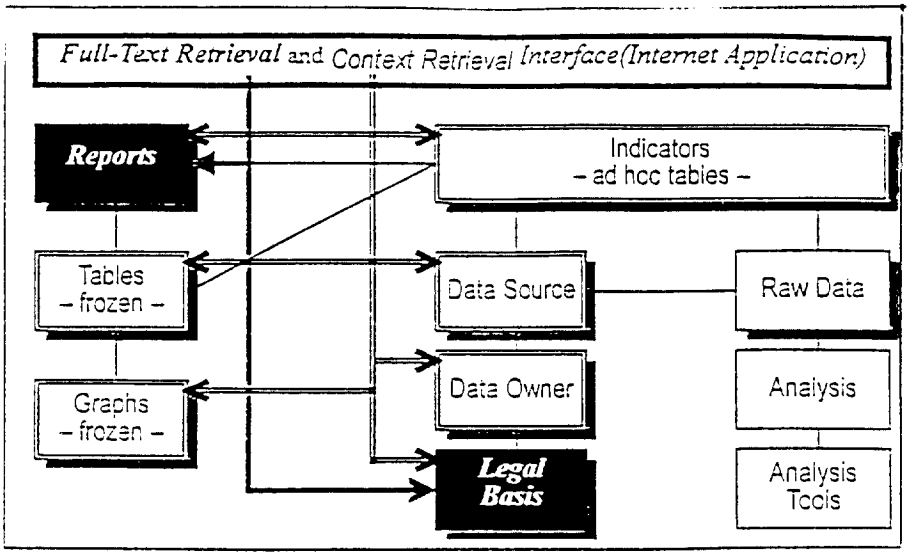
- La diversité des niveaux de connaissances, informatiques ou thématiques, *des utilisateurs* représente une véritable gageure pour les concepteurs d'un quelconque système d'information. La plupart posent des questions floues du fait, par exemple, d'une formulation imprécise. Pratiquement personne n'est familiarisé avec la hiérarchie des programmes parce qu'il n'existe pas de hiérarchie formelle commune et que la plupart des éléments d'information se rapportent à plus d'un fichier thématique. Les utilisateurs utilisent des langages différents, qu'ils manient avec plus ou moins de facilité.

- La communication d'informations numériques est un exercice particulièrement difficile du fait que les chiffres ne sont jamais explicites *per se* et n'ont de sens que dans un contexte très précis. Ce contexte doit être bien compris de l'utilisateur mais aussi du fournisseur si l'on veut éviter les erreurs d'interprétation et de fausses conclusions. C'est là que les indicateurs jouent un rôle important : ils donnent des réponses à des requêtes dont la formulation a été convenue par des experts. Le rôle des indicateurs est d'indiquer, mais ils sont inutiles sans indicandum, lequel est indispensable pour guider l'utilisateur dans le labyrinthe du système d'information.

III.2 Modélisation des données

10. La modélisation des données se divise en trois grandes parties (voir fig. 1). L'utilisateur communique avec l'application qui sert d'interface pour toutes les informations stockées. Qui dit recherche d'une information dit récupération de texte.

Figure 1 : Modélisation du système d'information sanitaire de l'Allemagne



Objets accessibles

11. Tous les objets encadrés d'un double trait dans la figure 1 peuvent être directement accessibles par l'utilisateur. Ceux encadrés d'un trait unique, dont "Raw Data" (données brutes) ou "Analysis Tools" (outils d'analyse) indiquent des programmes d'arrière-plan auxquels l'utilisateur n'a pas un accès prioritaire.

12. *Ces programmes* existent pour tous les objets ombrés de la figure 1. Ils contiennent des données d'aide ou des métadonnées. Ils permettent de comprendre ou de "lire" correctement les données. C'est ainsi qu'un programme d'arrière-plan relatif à des indicateurs peut, par exemple, donner la formule ou indiquer les domaines spécifiques d'application d'un indicateur. Il peut aussi mettre l'utilisateur en garde quant aux conclusions qui ne devraient pas être tirées de ses valeurs. Un programme d'arrière-plan sur des sources de données indiquera, par exemple, la périodicité, la taille de l'échantillon ou la liste des variables d'une acquisition de données, ce qui est particulièrement utile si des définitions ont changé lors de saisies régulières ou si des variables ont été ajoutées ou abandonnées. Le programme d'arrière-plan sur les propriétaires de données indiquera le nom et l'adresse de l'organisme qui a collecté les données ainsi que le nom de la personne à contacter. De tels renseignements sont utiles lorsqu'une source de données contient des variables qui intéressent l'utilisateur mais qui ne font pas partie intégrante du système d'information sanitaire.

Recherche de l'information

13. La recherche de l'information à partir des objets accessibles peut suivre deux cheminements : la recherche textuelle ordinaire ("context retrieval", fig. 1) utilise des index textuels qui sont automatiquement créés lorsqu'un objet est chargé dans la base de données, opération pendant laquelle tous les paramètres descriptifs associés à l'objet sont vérifiés et tous les mots indexés, sauf ceux explicitement conservés dans la "liste des codes d'arrêt". Cette forme de recherche textuelle est rapide et accessible pour tous les objets.

14. La recherche du texte intégral est une forme de consultation plus complexe, installée seulement pour les deux objets qui apparaissent en noir dans la figure 1 : "Reports" (Rapports) et "Legal basis" (Base juridique). Les utilisateurs peuvent opter pour une méthode de localisation plus pointue, celle de la chaîne par exemple des mots adjacents dans une phrase ou dans un paragraphe. Ils peuvent aussi rechercher des objets contenant une certaine chaîne de texte à l'exclusion d'une autre.

Catégories d'information

15. Tous les objets d'information sont groupés en deux grandes catégories : "Reports" (Rapports), "Tables-frozen" et Graphs-frozen" (Tables et graphiques figés) forment un groupe (colonne de gauche, fig. 1), ils se rapportent à l'analyse et à l'interprétation des données sous diverses formes. Les tableaux et les graphiques dit "figés" sont ceux qui ne sont pas automatiquement mis à jour lorsque de nouvelles données sont saisies, ce qui en garantit

l'authenticité lorsque ces tableaux ou ces graphiques ont déjà été publiés dans des rapports.

16. Dans la deuxième catégorie entrent les objets renvoyant à des données numériques. L'indicateur est la fenêtre de base de ces données; les données brutes ne sont pas accessibles à l'utilisateur si elles ne sont pas explicitement définies en tant qu'indicateurs. Des informations sont fournies sur les sources de données utilisées pour le calcul de l'indicateur, sur les propriétaires des données et sur la base juridique des saisies.

III.3 Technologie logicielle

17. Dans la mesure du possible, on a fait appel aux logiciels existants pour rester à une petite échelle d'application et faciliter la maintenance du système d'information sanitaire, en grande partie à des bases de données communes, dans notre cas un système de gestion de base de données relationnelles (SGBDR). Le calcul des indicateurs nous a posé plus de difficultés dues au fait qu'un SGBDR ordinaire n'a pas la souplesse nécessaire, mais fort heureusement, une technologie nouvelle est arrivée juste à temps, à savoir le traitement analytique en ligne (TAL).

18. Le TAL traite les données en tant qu'ensemble dimensionnel -n, les dimensions étant identiques aux catégories de répartition, d'où son appellation de système de gestion de bases de données multidimensionnelles (SGBDM). Les données relatives à la mortalité, par exemple, sont le type même des données saisies et communiquées annuellement, ventilées par âge, sexe, région et cause (code ICD - input command data). TAL a donc organisé les données de mortalité en cinq ensembles dimensionnels selon les quatre catégories mentionnées, plus un paramètre temps. Pour toutes les dimensions, des niveaux de répartition hiérarchique multiples peuvent être définis : selon l'âge, par exemple, des groupes d'âges de cinq ans associés à des groupes de 10 ans et une catégorie "moins de 65 ans" et une autre "plus de 65 ans". Au moment de la saisie des données dans le SGBDM, tous les sous-totaux pour les positions définies de la répartition sont automatiquement établis.

19. Le calcul des indicateurs implique désormais l'application à ces ensembles de matrices algébriques. Les résultats sont automatiquement calculés et présentés sous forme de tables sur la base d'un "aperçu type" prédéfini. La souplesse du système est assurée par deux fonctions : *matriçage et permutation*. Le matriçage ouvre un niveau hiérarchique plus détaillé en partant du bas pour arriver aux valeurs maximales définies pour la série de données. La permutation permet d'arranger les dimensions de toutes les manières possibles, et, partant, de déterminer la disposition spécifique de la table, en rangs et colonnes.

20. Les valeurs des indicateurs ne sont pas stockées dans la base de données mais calculées uniquement à la demande de l'utilisateur. Avec TAL, calculer un indicateur, au passage, ne prend pas plus de temps qu'il ne faudrait pour le rechercher dans le système. Le stockage de valeurs précalculées grossirait énormément la base de données sans grand avantage. Seules des données brutes sont stockées dans le SGBDR et chargées dans le SGBDM. Le contrôle

d'authenticité ne s'applique qu'au SGBDR. Des indicateurs peuvent être librement ajoutés en choisissant une formule et en précisant la marque de réservation; ces indicateurs seront alors accessibles à tout le monde. Il peut être intéressant de savoir que le système d'information sanitaire contient plus de 1 700 indicateurs ayant pour base moins de 50 formules.

III.4 Matériels et logiciels utilisés

21. La continuité requise a présidé au choix de la plate-forme matérielle et des fournisseurs de logiciels. Nous avons choisi une plate-forme UNIX et des produits ORACLE. La configuration du système est la suivante :

Matériels

22. Le SGBDR et le serveur TAL sont installés sur un serveur parallèle jumeau SUN-Enterprise 3000. Ce système est protégé par un pare-feu, contre Internet et Intranet. Un serveur Web sert de messager, échange des informations entre l'utilisateur et le système au travers du pare-feu. Les matériels installés sont les suivants :

- 2 SUN Enterprise 3000 utilisant une pile de disques de 25 GB;
- 1 SUN Ultra Enterprise 2 servant de pare-feu Internet;
- 1 SUN Ultra Enterprise 2 contenant le serveur Web pour Internet;
- 1 SUN Sparcstation 5 agissant en tant que pare-feu Intranet;
- 1 SUN Sparcstation 20 contenant le serveur Web pour l'Intranet;
- SUN OS 5.5.1. pour toutes les plates-formes

Logiciels

23. Toutes les informations du système sont stockées dans un SGBDR servant de module de traitement transactionnel en ligne (MTTL). TAL, la recherche textuelle et la connexion avec Internet exigent des éléments supplémentaires. Les logiciels suivants sont installés :

- ORACLE 7.3, serveur parallèle SGBDR;
- ORACLE Express Objects 6.0, serveur TAL;
- ORACLE Context 2.0, élément de recherche textuelle;
- ORACLE Web Server 3.0.x, serveur Web.

Logiciels nécessaires à l'utilisateur

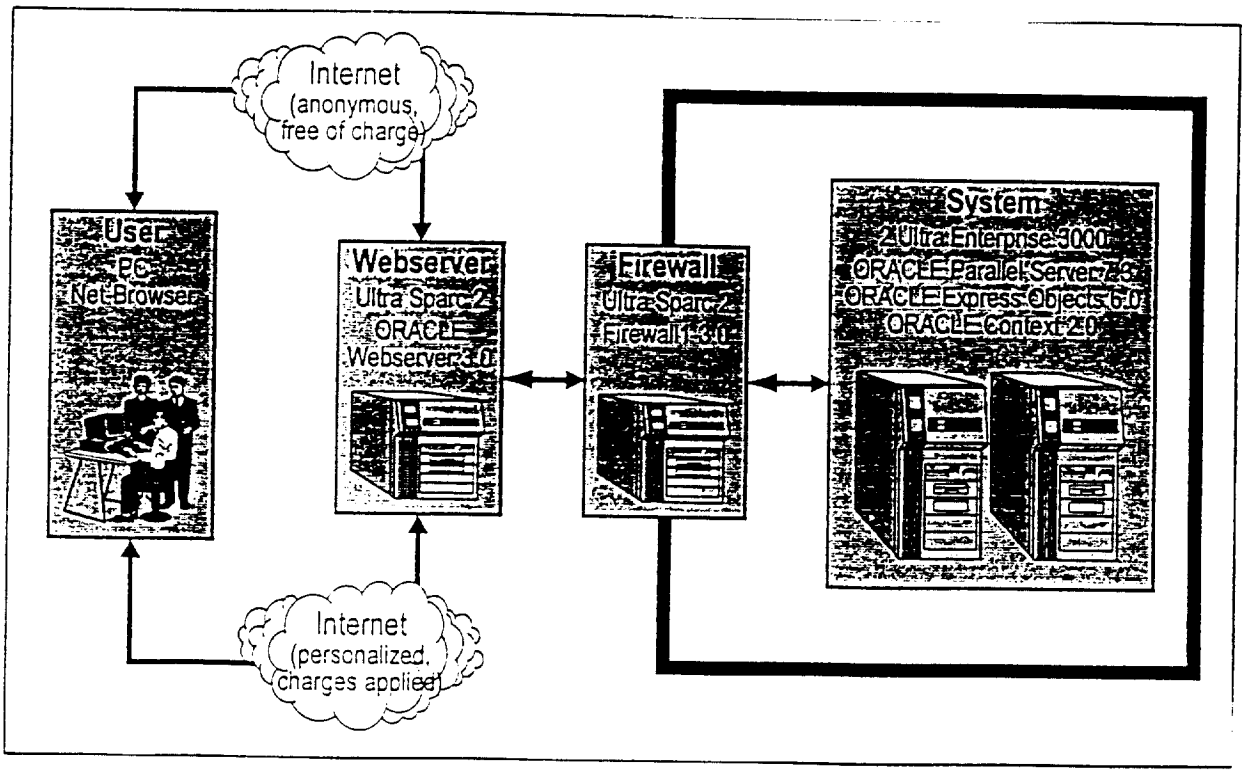
24. L'utilisateur n'a besoin que d'un navigateur Internet ordinaire, tel que Netscape Navigator ou Microsoft Internet Explorer, indépendamment d'un accès à Internet. S'il veut procéder à un post-traitement ou analyser l'information communiquée par le système sur sa propre plate-forme, il aura besoin d'autres

logiciels, dont un traitement de textes, un programme tableur ou un système d'analyse statistique.

III.5 Accès au système

25. La figure 2 montre les éléments actifs lors d'une connexion classique avec Internet. Chacun peut choisir d'utiliser le système anonymement (sans frais) ou de manière personnalisée (à titre onéreux) pour avoir plus de choix.

Figure 2 : Accès au système d'information sanitaire



III.6 Normes ergonomiques en matière de logiciels

26. Depuis le 1er janvier 1998, tous les États membres de l'Union européenne doivent satisfaire à des normes ergonomiques en matière de conception de logiciels. Le Système d'information sanitaire a été le premier logiciel de l'Office fédéral des statistiques à être testé compte tenu de ces normes. Des conseillers extérieurs spécialisés ont aidé l'équipe de conception à cet égard. Des tests pratiques ont été effectués par des employés de l'Office statistique qui n'avaient pas participé à la conception des logiciels et un "test bêta" a été réalisé auprès d'utilisateurs externes. Les résultats de ces deux séries de tests sont exploités par les concepteurs pour l'amélioration du Système.

IV. Utilisation du Système d'information sanitaire

27. Les concepteurs se sont ingéniés à rendre le Système d'information sanitaire facile à utiliser. Une aide en ligne sera mise à la disposition des personnes intéressées, et notamment une démonstration de recherche. Tout internaute familiarisé avec un moteur de recherche sur Internet (Altavista ou Yahoo) trouvera des similitudes, ce qui n'est certainement pas le fait du hasard.

28. Il est possible d'accéder au Système en utilisant n'importe quel navigateur Internet ordinaire. Même les versions plus anciennes qui ne sont pas compatibles avec les trames de langage pour la création de liens hypertextes (HTML) sont prises en charge. L'adresse Internet est la suivante :

HTTP://WWW.GBE-BUND.DE

29. "GBE" est l'acronyme allemand pour surveillance sanitaire (Gesundheitsberichterstattung), "Bund" indique que le Système est concentré au plus haut niveau régional et qu'on n'y trouvera pas de données relatives aux comtés ou aux municipalités. La décision de limiter le Système d'information sanitaire à deux niveaux, le territoire fédéral et les Länder, découle du partage traditionnel des tâches entre les administrations allemandes et non de considérations techniques.

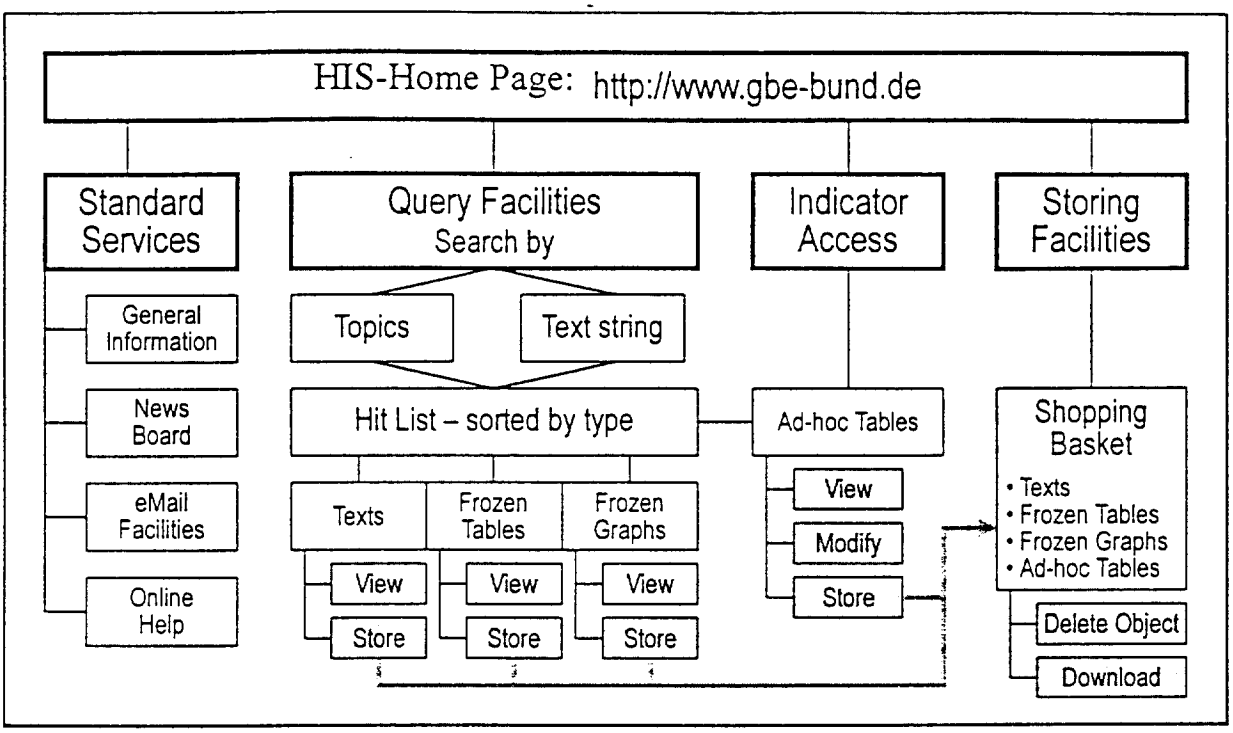
IV.1 Services fournis

30. Le Système d'information sanitaire donne à l'utilisateur tous les moyens nécessaires à la recherche d'informations et à leur importation sur son propre ordinateur, pour plus ample traitement s'il le souhaite.

31. Les services dont l'utilisation n'est pas continue, tels que "General information on the system", "News board" ou "Communication with the Administrator and the team", sont concentrés sur la page d'accueil qui est le seul point d'accès, d'autres, par exemple les outils de navigation, sont disponibles pendant toute la session parce qu'ils font partie intégrante de chaque image-écran.

32. La figure 3 montre l'organigramme de tous les services. Les types de services sont indiqués par des cases encadrées d'un trait gras et les activités des utilisateurs par des cases ombrées.

Figure 3 : Organigramme des services offerts à l'utilisateur



IV.2 Page d'accueil du Système d'information sanitaire

33. Le Système est accessible en deux langues, allemand et anglais. La version allemande apparaît par défaut. Pour atteindre la version anglaise il faut activer l'icône représentant le drapeau britannique sur la page d'accueil. Le résultat de cette activation est montré à la figure 4.

34. Les six éléments qui figurent en haut de la page servent d'outils de navigation et apparaissent sur toutes les autres pages. Lors d'une session, on peut passer de l'une à l'autre de ces options à tout moment et aussi souvent que voulu. Il vaut mieux éviter d'utiliser les fonctions de balayage intégrées "Back" ou "Home" qui peuvent provoquer des dysfonctionnements.

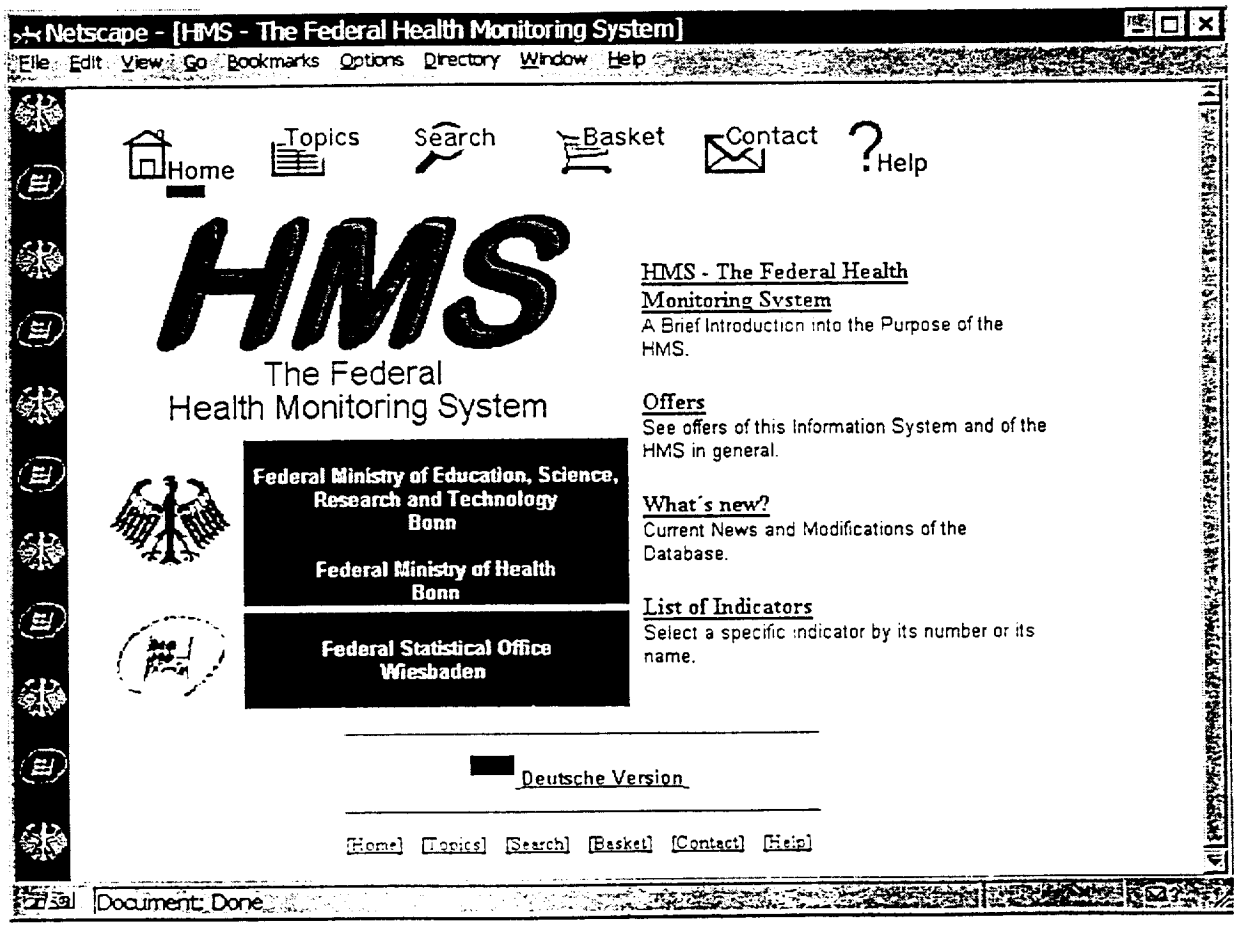
35. La position de navigation "active" est marquée par un trait de soulignement dans une couleur spéciale qui réapparaît dans la bande verticale sur le côté gauche de l'écran. La couleur est verte pour la page d'accueil (voir fig. 4).

36. Les outils de navigation offrent trois options qui s'expliquent d'elles-mêmes. "Contact" permet à l'utilisateur de communiquer avec l'équipe du Système d'information sanitaire par courrier électronique (voir "email Facilities", fig. 3), "Help" active l'aide en ligne et "Home" renvoie l'utilisateur à la page d'accueil.

37. Les trois autres outils de navigation permettent à l'utilisateur de choisir l'une ou l'autre des techniques de récupération - recherche par thème ("Topics") et recherche par chaînes de texte ("Search") - et d'entreposer les objets sélectionnés dans une corbeille personnelle ("Basket") pour importation ultérieure.

38. Entre autres options offertes par la page d'accueil, dont "News Board", figurent celles qui permettent à l'utilisateur averti de choisir directement des indicateurs sur une liste comparable à la version en ligne de la base de données SPT du Bureau régional de l'OMS pour l'Europe (voir chap. 4.6 "The indicator page - OLAP facilities" pour plus de détails). Cette option apparaît dans la case "Indicator Access" à la figure 3.

Figure 4 : La page d'accueil du Système d'information sanitaire



IV.3 Recherche par thème

39. Le programme "Search by topics" donne un aperçu de la structure thématique du Système d'information sanitaire. Cette structure suit les sept principales catégories du "Rapport sanitaire pour l'Allemagne" et les 32 sous-catégories s'y rapportant. Le nombre des éléments détaillés ne manquera pas d'augmenter à l'avenir mais la liste des thèmes et des

sous-thèmes a été définie de manière à couvrir d'ores et déjà tout l'éventail des informations relatives à la santé dont le traitement est prévu par le Système. Cette liste devrait donc rester inchangée.

40. La recherche par thème est une stratégie particulièrement utile pour les débutants ou les utilisateurs dont les recherches sont moins spécifiques. À mesure que le Système se développera, cette stratégie deviendra, à plus ou moins long terme, moins attrayante. Un taux excessif de présence en antémémoire rendra le survol des objets montrés difficile, même si le Système peut établir des priorités en classant les objets de manière à montrer en premier ceux ayant le moins d'antériorité.

41. Lorsqu'un thème et un sous-thème sont sélectionnés, la possibilité est donnée à l'utilisateur de réduire le champ de recherche selon ses besoins ou de consulter l'ensemble du fichier du sous-thème "Entire subtopic". Cette option peut être la plus judicieuse tant que le Système reste relativement petit, mais elle sera à déconseiller à l'avenir lorsque le Système se sera considérablement développé.

Figure 5 : La page "Search by topics" (recherche thématique)

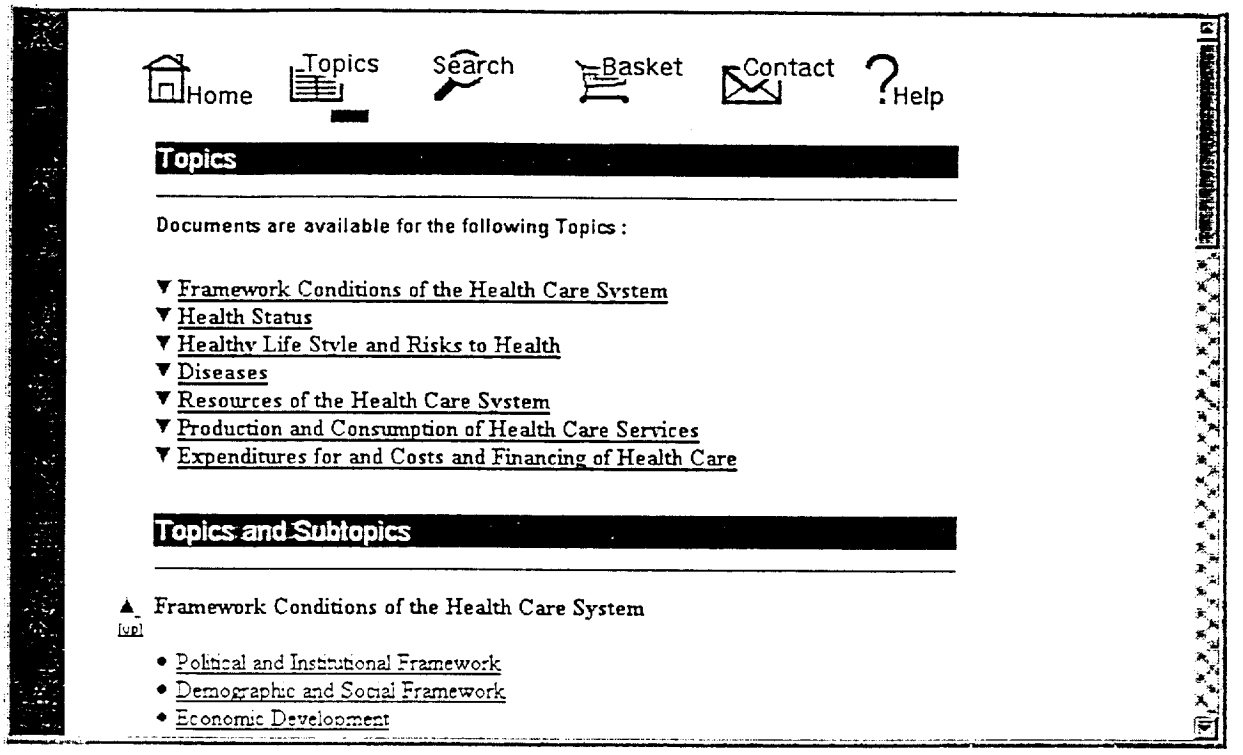
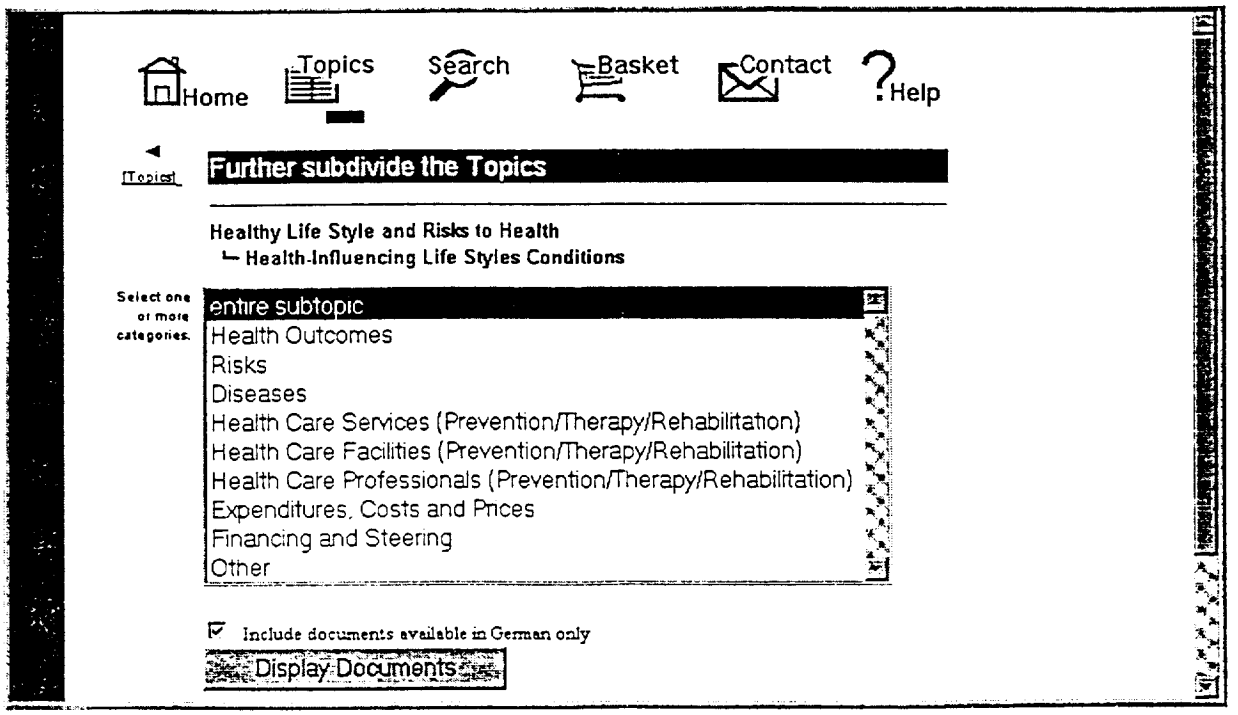


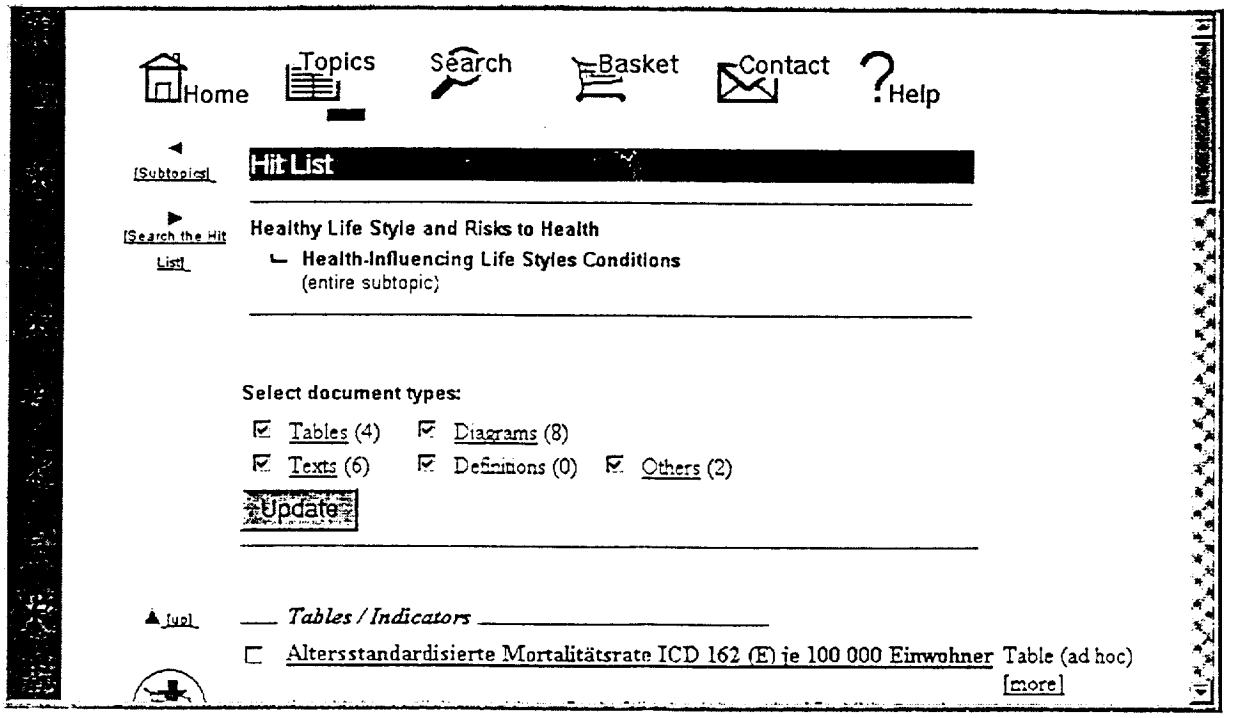
Figure 6 : La page "Display documents" (affichage des documents)



42. La case "Include documents available in German only" appelle quelques observations. Le Système d'information sanitaire ayant été conçu avant tout pour l'Allemagne, la langue principale est l'allemand. Nous avons l'intention de fournir une version bilingue des descriptions de base, mais seulement lorsque nous aurons atteint notre objectif prioritaire qui est de parachever le Système pour le rendre disponible dans la langue de référence. Nous serons alors en mesure de traduire toutes les descriptions et de fournir des résumés en anglais de tous les rapports, mais nous ne pourrons jamais assurer leur traduction *in extenso*, nos possibilités de traduction étant insuffisantes. Aussi proposons-nous la solution suivante, espérant qu'elle donnera satisfaction : si un utilisateur n'a pas une connaissance suffisante de l'allemand pour comprendre un texte donné mais pense, d'après le résumé, qu'il pourrait lui être utile, il peut le télécharger et le faire traduire lui-même.

43. Il est donc recommandé d'activer la case "Include documents available in German only", faute de quoi des objets pouvant présenter un certain intérêt ne seront pas affichés.

Figure 7 : Affichage des objets localisés

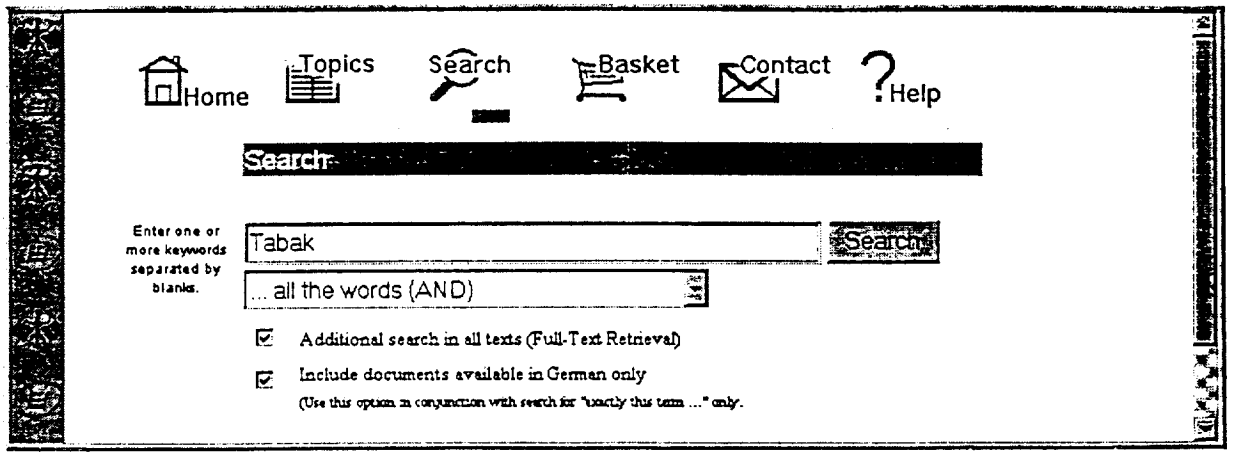


44. L'affichage des objets localisés commence par un inventaire, par type de documents. Cliquer une catégorie permet d'obtenir la description du document; cliquer le titre d'un document permet d'accéder au texte; cliquer l'enchaînement "More" permet d'obtenir une description plus complète ou un extrait du document. L'utilisateur peut envoyer directement l'objet dans sa corbeille de récupération ("Basket") sans l'avoir d'abord survolé en activant la case correspondante et en pressant le symbole "Put into basket" sur le côté gauche de la fenêtre. Lorsqu'une opération de "Shopping" (Récupération) est fructueuse, la page se déroule et le symbole "Put into basket" est temporairement remplacé par le symbole "OK".

IV.4 Recherche par chaînes textuelles

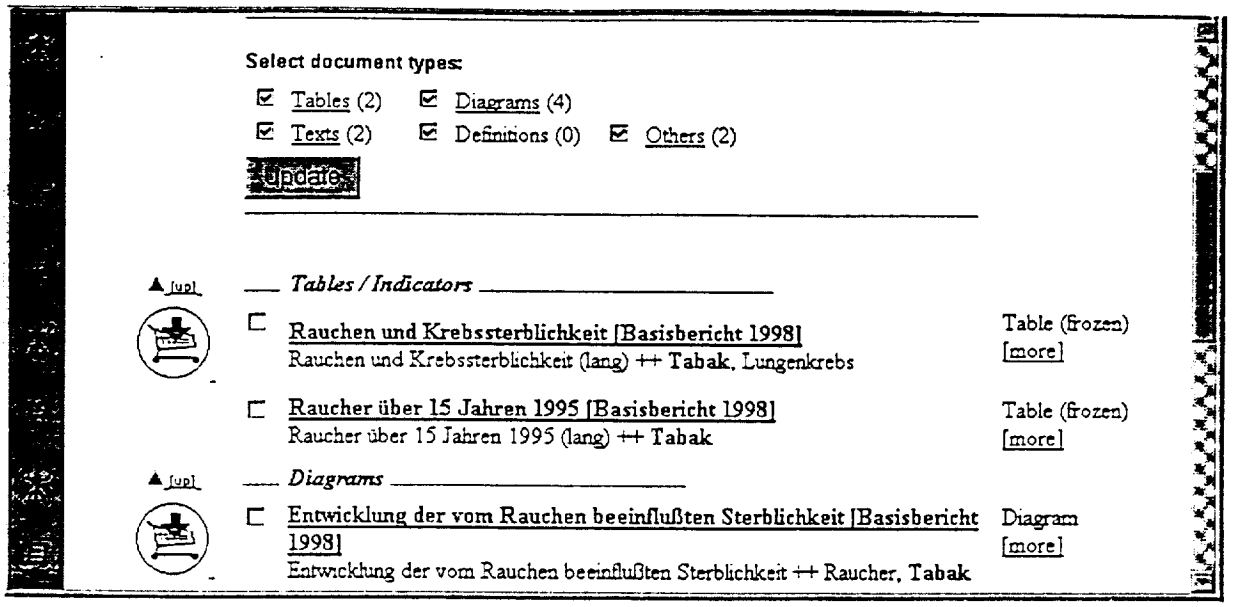
45. Presser le symbole de navigation "Search" donne accès au service "Search by text strings". Les chaînes textuelles introduites avec les options "... all the words (AND)" et "... any of the words (OR)" feront apparaître les objets contenant toutes les chaînes textuelles, simultanément, ou au moins l'une d'elles, respectivement. Les autres options sont "... exactly this term ("...")" (... exactement ce terme), "... exactly this term and its synonyms" (... exactement ce terme et ses synonymes) et "... exactly this term and similar sounding terms" (... exactement ce terme et les termes homophones).

Figure 8 : La page "Search by text strings" (Recherche par chaînes textuelles)



Introduire le mot "Tabak" (Tabac) fera apparaître l'inventaire y relatif avec la chaîne de recherche imprimée en gras.

Figure 9 : Affichage des résultats de la recherche par chaînes textuelles

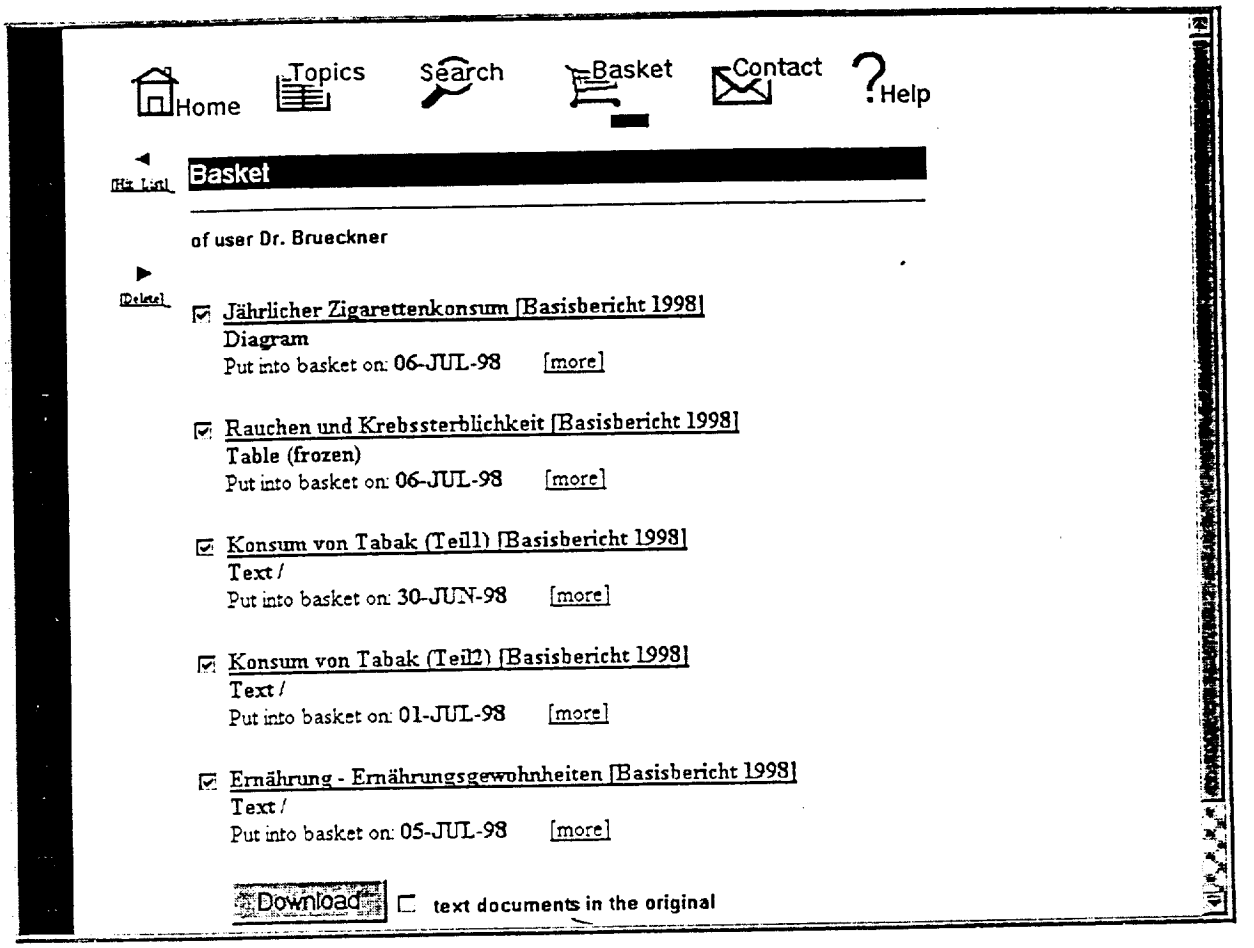


La suite de la recherche à partir de là est la même que pour "Search by topic" décrite ci-dessus.

IV.5 Corbeille de récupération (Shopping basket)

46. La corbeille de récupération contient tous les objets qu'un utilisateur anonyme y a mis pendant une session de recherche. Pour l'utilisateur personnalisé elle pourra contenir des objets récupérés pendant des sessions antérieures, dans la mesure où ces objets n'auront pas été déjà téléchargés ou annulés.

Figure 10 : La page "Basket" (Corbeille)



47. La corbeille de récupération peut être partiellement ou entièrement vidée en choisissant l'option "Delete" (Supprimer) pour les documents dont la case de contrôle aura été cliquée. L'option "Download" (Télécharge) assurera le transfert des documents choisis sous forme compressée. L'utilisateur doit indiquer s'il utilise une plate-forme INTEL ou UNIX. Les textes et tables figés ont le code HTML; ils peuvent être post-traités, avec un MS-Word 97, par exemple. Le format d'échange graphique (GIF) est celui utilisé pour les graphiques. Le code des tables ad hoc est le code standard américain d'échange d'informations (ASCII), aussi peuvent-elles être chargées dans n'importe quel tableur.

48. La case "Text documents in the original" une fois cliquée, les textes activés seront téléchargés sous le format PDF (Portable display format). Le style de caractères et la disposition de la page correspondront exactement à la version imprimée. Le PDF ne se prêtant pas, en général, à un traitement plus complexe, l'utilisateur voudra peut-être télécharger un texte à la fois en HTML et en PDF.

IV.6 Page des indicateurs - Services TAL (Traitement analytique en ligne)

49. Les indicateurs représentent le service le plus novateur du Système d'information sanitaire. On peut y accéder soit par nom, via le lien correspondant sur la page d'accueil (voir chap. 4.2) ou, le plus souvent, via des documents étiquetés "ad-hoc tables" dans l'inventaire. "Life expectancy" (Espérance de vie) est subséquemment utilisé en tant qu'indicateur exemplaire parce qu'il incorpore un mode de calcul plus pointu. Avant l'affichage de la table, une version abrégée a été créé de a à z, selon la méthode Farr.

Figure 11 : Tabulation ad hoc de l'espérance de vie

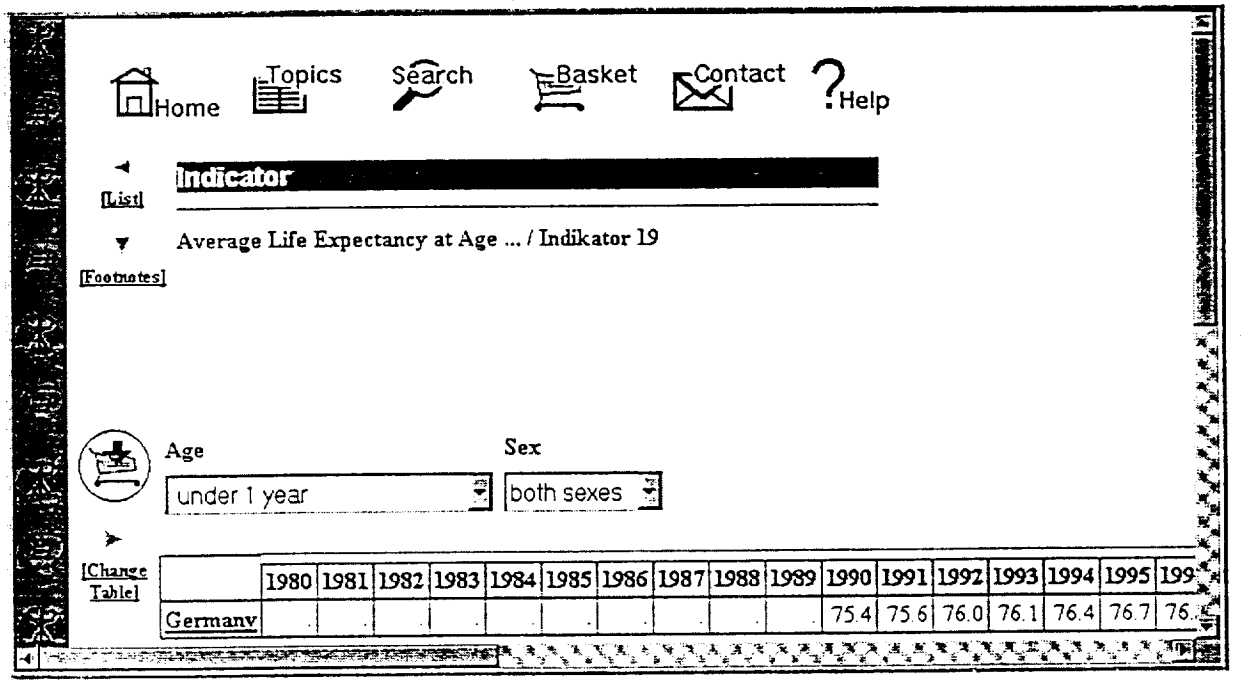


Figure 12 : Réaction TAL à "Drill-down" (Déroulement) et "Change page" (Changement de page)

Age: under 1 year — Sex: female

[Change Table]

	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
Germany	.	.	.	.	.	.	.	.	.	.	78.4	78.7
Schleswig-Holstein	76.8	76.9	77.3	77.6	77.9	78.3	78.5	78.6	78.7	79.0	78.9	79.1
Hamburg	76.4	76.6	76.6	77.3	77.8	77.8	77.8	78.3	78.9	78.9	79.2	78.9
Lower Saxony	76.8	76.9	77.1	77.6	78.0	78.2	78.3	78.7	78.9	79.0	79.0	79.2
Bremen	76.8	76.5	77.5	77.3	78.1	78.2	77.8	78.5	78.3	79.1	78.8	78.9
North Rhine-Westfalia	76.2	76.5	76.8	77.0	77.7	77.8	77.9	78.5	78.7	78.7	78.9	79.1
Hesse	77.1	77.1	77.5	77.7	78.2	78.3	78.6	79.0	79.0	79.1	79.1	79.3
Rhineland-Palatinate	76.5	76.4	76.9	77.0	77.4	77.9	78.0	78.5	78.8	78.9	78.8	79.2
Baden-Württemberg	77.5	77.6	78.0	78.2	78.9	78.9	79.3	79.6	79.8	79.8	79.7	80.1
Bavaria	76.7	77.0	77.3	77.4	78.0	78.2	78.6	78.7	79.0	79.0	79.1	79.4
Saarland	75.5	75.8	76.2	76.2	76.9	76.8	77.2	78.0	78.1	78.1	78.1	78.5
Berlin	69.0	69.5	69.7	70.2	70.3	70.6	71.0	72.1	72.2	72.8	77.1	77.2
Brandenburg	.	.	.	.	.	.	.	.	.	.	76.1	76.4
Mecklenburg-Western Pomerania	.	.	.	.	.	.	.	.	.	.	75.6	76.0
Saxony	.	.	.	.	.	.	.	.	.	.	76.7	77.1
Saxony-Anhalt	.	.	.	.	.	.	.	.	.	.	76.0	76.2
Thuringia	.	.	.	.	.	.	.	.	.	.	76.1	76.9

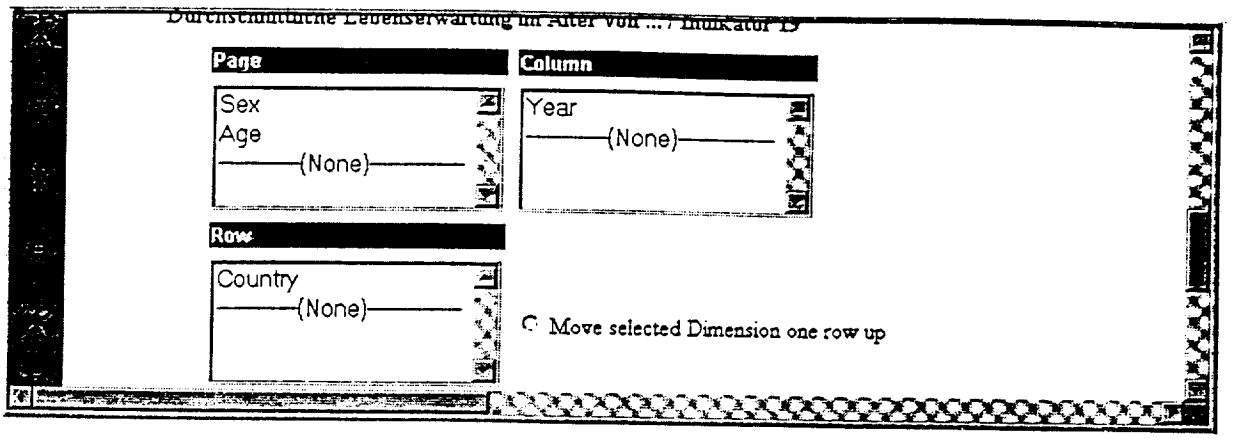
50. La présentation par défaut d'un indicateur se rapporte à l'Allemagne et le total aux deux sexes. En faisant défiler le menu, l'utilisateur peut choisir "Female"; s'il clique "Germany", il ouvre le niveau des Länder. La figure 12 montre la réaction TAL.

51. Les valeurs "Espérance de vie" manquantes avant 1990 sont indiquées pour les nouveaux Länder à la figure 12. Ces lacunes sont dues au fait que les moyennes démographiques ne sont pas encore disponibles sous une forme adaptée aux limites territoriales des nouveaux Länder. TAL insère automatiquement le tiret pour indiquer les valeurs manquantes dans l'une des sources de données utilisées pour le calcul.

IV.7 Autres services TAL mis à la disposition des utilisateurs personnalisés

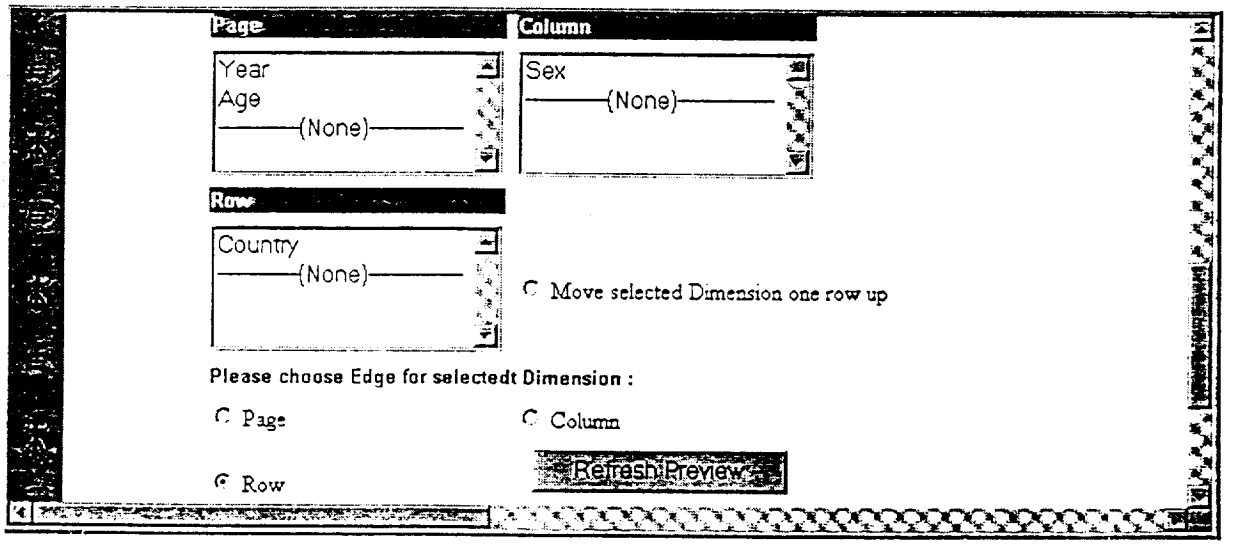
52. Les utilisateurs personnalisés ont accès à l'option "Change table" qui leur permet de permuter la présentation d'une table pour l'adapter à leurs besoins. Le menu représenté à la figure 13 est celui utilisé.

Figure 13 : Modification de la présentation de la table de l'indicateur "Life expectancy" (Espérance de vie)



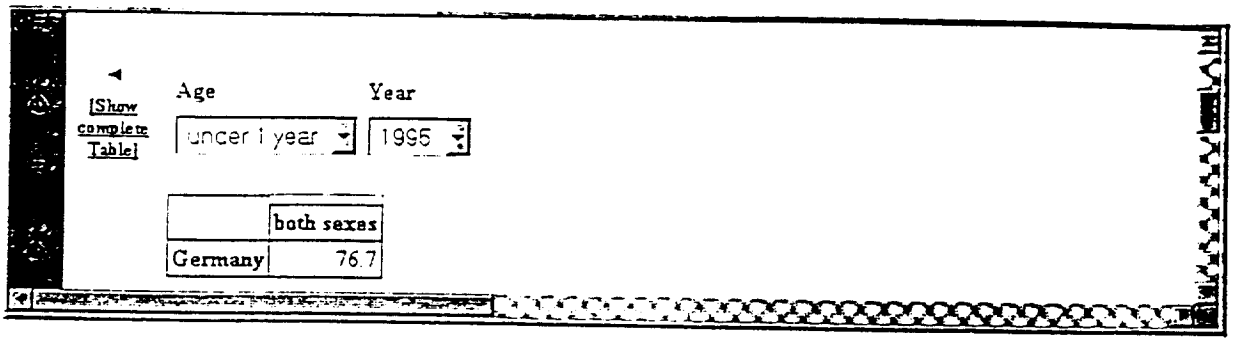
53. Le menu montre que le sexe et l'âge sont attribués à "Page", l'année à "Column" et le pays à "Row". Si des indicateurs d'espérance de vie devaient être affichés pour la seule année 1995, avec les régions par rang et le sexe par colonne, il serait nécessaire de faire permuter les éléments de la table.

Figure 14 : Permutation des éléments de la table



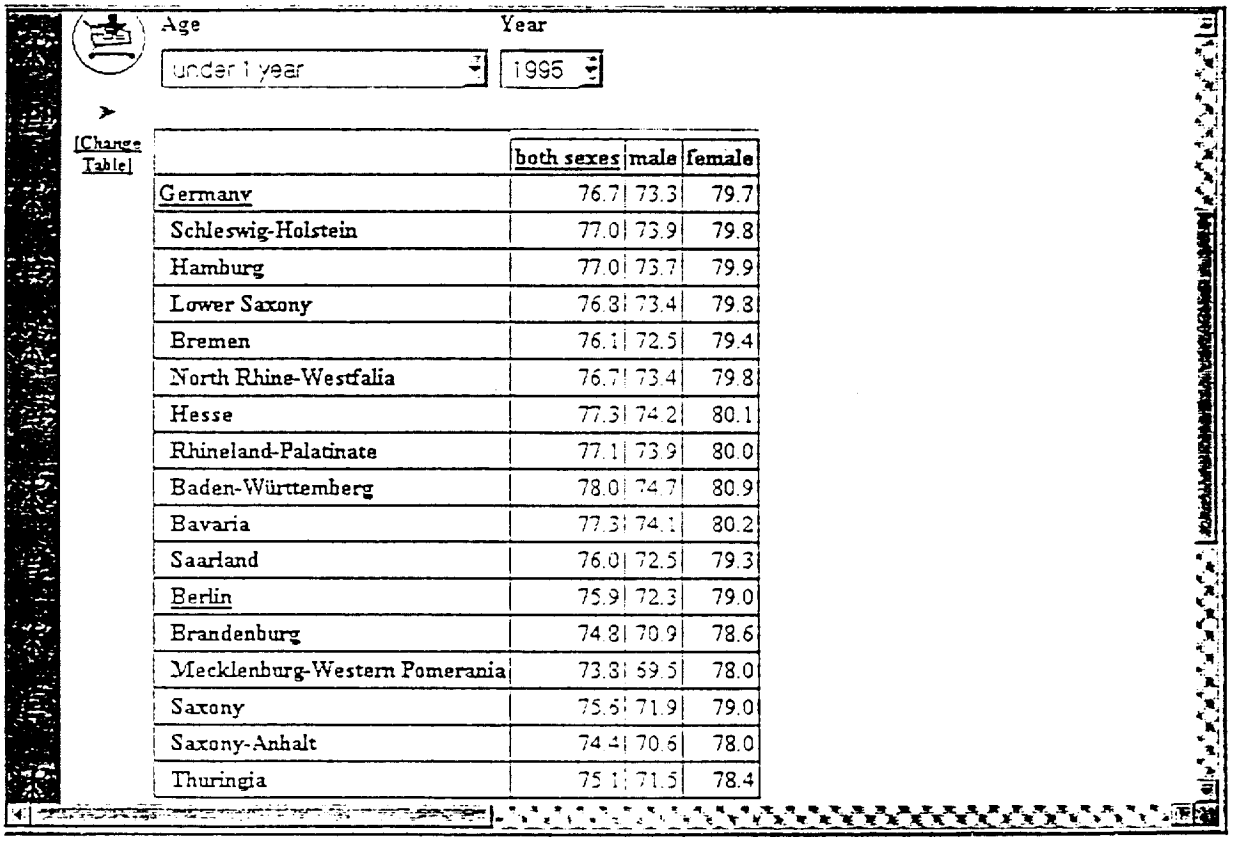
54. "Sex" devrait être permuté sous "Column" et "Year" sous "Page". Après cette opération, l'année 1995 peut être choisie comme référence. Pour procéder à une permutation, l'utilisateur doit entrer une dimension, contrôler les boutons correspondants et cliquer sur "Refresh preview", comme indiqué à la figure 14. Cette procédure devra être répétée pour chaque dimension à déplacer. La table ainsi modifiée pourra alors être affichée par le lien "Show complete table" (voir fig. 15). Enfin, il faudra choisir "Année 1995".

Figure 15 : Activation de l'affichage de la table après permutation



Ensuite, TAL affichera la table telle que modifiée pour l'espérance de vie à la naissance, par région et par sexe (voir fig. 16).

Figure 16 : Affichage de la table après permutation



	both sexes	male	female
Germany	76.7	73.3	79.7
Schleswig-Holstein	77.0	73.9	79.8
Hamburg	77.0	73.7	79.9
Lower Saxony	76.8	73.4	79.8
Bremen	76.1	72.5	79.4
North Rhine-Westfalia	76.7	73.4	79.8
Hesse	77.3	74.2	80.1
Rhineland-Palatinate	77.1	73.9	80.0
Baden-Württemberg	78.0	74.7	80.9
Bavaria	77.3	74.1	80.2
Saarland	76.0	72.5	79.3
Berlin	75.9	72.3	79.0
Brandenburg	74.8	70.9	78.6
Mecklenburg-Western Pomerania	73.8	69.5	78.0
Saxony	75.6	71.9	79.0
Saxony-Anhalt	74.4	70.6	78.0
Thuringia	75.1	71.5	78.4

Un utilisateur qui déciderait de placer cette table dans la corbeille de récupération n'y trouvera que les parties visibles après téléchargement. S'il veut télécharger des informations pour d'autres années, il lui faudra

répéter la procédure mentionnée ci-dessus pour chaque année et récupérer les résultats dans la corbeille, année après année.

Résumé

55. La conception du Système d'information sanitaire est le fruit des efforts conjugués de "Robotron Data Systems in Dresden" (société de développement), de l'Office fédéral des statistiques (client) et du "Bureau for Applied Organizational Psychology" (ergonomie logicielle). Au total, l'installation du système a pris 12 mois. Jusqu'à sept collaborateurs de Robotron ont fait partie de l'équipe, plus trois de l'Office des statistiques, soit au total dix années-homme.

56. Le coût total du système est élevé, soit 4,5 millions de DM environ. Le développement de l'application a absorbé quelque 60 % de ce montant et le reste a servi à l'achat des matériels et des licences de logiciels. Quoi qu'il en soit, la décision de concevoir un système d'information sur Internet utilisant des logiciels de bases de données normalisés et les outils TAL s'est révélée très rentable. Mettre en place un tel système de a à z aurait exigé beaucoup plus de ressources, aussi bien en temps qu'en argent.