

Réalisation de la base FINES (Fichier national de l'Expertise Scientifique), un répertoire d'experts des grands organismes de recherche français

Pierre Pélisson

▶ **To cite this version:**

Pierre Pélisson. Réalisation de la base FINES (Fichier national de l'Expertise Scientifique), un répertoire d'experts des grands organismes de recherche français. [Rapport de recherche] Institut de l'Information Scientifique et Technique (INIST-CNRS). 1996, 72 p. hal-01457454

HAL Id: hal-01457454

https://hal-lara.archives-ouvertes.fr/hal-01457454

Submitted on 6 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INIST

CONTRAT MENESR FINES 94K5226

Rapport de fin de contrat

Réalisation de la base expérimentale FINES (Fichier National de l'Expertise Scientifique), un répertoire d'experts des grands organismes de recherche français

auteur : Pierre-Gilles Pélisson

8 juillet 1996

CONTRAT MENESR FINES 94K5226

Rapport de fin de contrat

I. RAPPELS
- **I.1. Objet du contrat.**
- **I.2. Présentation de l'INIST.**
- **I.3. Contexte de l'étude.**
- **I.4. Objectifs.**

II. HISTORIQUE DES REALISATIONS
- **II.1. Préparation du fichier global.**
- **II.2. Tests, chargement et mise en œuvre de la maquette.**

III. SYNTHESE DES EVALUATIONS

Annexes techniques
1. Format de la base, DTD.
2. Normalisation des fichiers.
3. Copies d'écrans.
4. Grilles d'évaluation.
5. Contenu de la maquette.
6. Bilan IDT 96.
7. Comptes rendus de réunions du Comité de Pilotage.
8. Arrêté au Journal Officiel.

I. RAPPELS

I.1. OBJET DU CONTRAT

A la demande de la Direction de l'Information Scientifique et Technique et des Bibliothèques du Ministère de l'Enseignement Supérieur et de la Recherche, l'INIST a été chargé de la réalisation de la première version de la base expérimentale FINES (Fichier National de l'Expertise Scientifique).

Cette opération comprend la constitution d'un fichier global rassemblant des données provenant de sept organismes, CEA, CNRS, CIRAD, INRA, INRIA, INSERM et ORSTOM, ainsi que la réalisation d'une maquette interrogeable en ligne, préfigurant la future base.

I.2. PRÉSENTATION DE L'INIST

L'Institut de l'Information Scientifique et Technique (INIST) du Centre National de la Recherche Scientifique (CNRS) est un centre de collecte, de traitement et de diffusion de l'information scientifique et technique. L'INIST exploite un fonds multidisciplinaire de 25 000 titres de périodiques et 195 000 rapports scientifiques, thèses, comptes-rendus de congrès, ouvrages. Il produit deux bases de données multidisciplinaires et multilingues : PASCAL en Sciences, Technologie, Médecine et FRANCIS en Sciences Humaines, Sciences Sociales et Economie.

L'INIST propose une gamme variée de produits et services : accès sur serveurs, Minitel ou CD-ROM à 13 millions de références bibliographiques, fourniture de la copie du texte des articles scientifiques et techniques disponibles dans son fonds ou auprès d'un réseau de bibliothèques européennes de recours, recherches bibliographiques sur demande, publications bibliographiques, traduction de tout document scientifique, conseil et formation, accès à l'information japonaise.

Le personnel de l'INIST (400 personnes) hautement qualifié est composé en majeure partie d'ingénieurs, experts en information spécialisée, bibliothécaires, documentalistes, informaticiens, dont les compétences seront mises à contribution dans ce projet.

I.3. CONTEXTE DE L'ETUDE

En juin 93, la Direction de l'Information Scientifique et Technique (DIST) du Ministère de l'Enseignement Supérieur et de la Recherche (MESR) décide de créer un Fichier National de l'Expertise Scientifique. Baptisée FINES, la future base permettra la collecte et la diffusion des fichiers d'experts des grands organismes de recherche français vers l'industrie et les décideurs politiques.

La décision du MESR intervient à un moment où la mise en oeuvre d'une base française de compétences scientifiques et techniques est devenue difficilement contournable : l'enjeu est à la fois scientifique et stratégique, économique et culturel.
Le projet de création d'un fichier informatique unique d'expertise, élaboré sous la responsabilité de grands organismes français apparaît à ces divers titres riche d'implications et constitue en outre une opportunité de réflexion sur l'application de travaux de recherche.

FINES se veut système unique et intégré des organismes français, avec un caractère institutionnel et systématique, à travers et sous la responsabilité des dits organismes.

I.4. OBJECTIFS

En permettant le repérage des experts (chercheurs, ingénieurs ou techniciens)par les industriels et en favorisant ainsi la mise en place de projets mixtes, FINES est essentiellement conçu de façon à faciliter le transfert des compétences des chercheurs et experts vers le monde industriel, dont les PME/PMI, éventuellement au travers d'interfaces comme, par exemple, les Centres Techniques Industriels.

La base doit également servir d'outil aux décideurs scientifiques, en permettant de cerner les zones où la recherche est active et importante.

Le recouvrement de l'expertise collective et de l'expertise individuelle devrait être facilité par la mise en place de ce système d'information globalisé et intégré. La communication interne dans le milieu scientifique, notamment dans le cadre de recherches multidisciplinaires, en sera améliorée.

Ainsi, FINES a pour objectif d'aider à la prise de décision au niveau des grandes questions de la politique de la recherche, tout comme de fournir des réponses à des interrogations précises et spécialisées.

Enfin, sur la scène internationale, FINES devra jouer le rôle de vitrine des compétences françaises à l'étranger et de réservoir d'experts, dans le domaine du développement.

II. HISTORIQUE DES REALISATIONS

La gestion du projet FINES a été confiée à un comité de pilotage qui s'est réuni pour la première fois le 18 juin 1993. Présidé par F. Hinard, Directeur de la Direction de l'Information Scientifique et Technique et des Bibliothèques du MESR, il rassemble les représentants des organismes participants (CEA, CIRAD, CNRS, INRA, INRIA, INSERM, ORSTOM et INIST).

Ce comité a orienté ces travaux sur trois points : d'une part, la gestion de la confidentialité. Un groupe de travail, placé sous la responsabilité du Secrétariat Général de la Défense Nationale (SGDN) a commencé ses travaux en novembre 93. Il est chargé de mettre en place une surveillance continue des données entrées dans FINES, sous l'observation de Hauts Fonctionnaires de la Défense.

Le Comité de Pilotage a d'autre part précisé les critères de prise en compte des individus dans la base.

Autres questions débattues : l'établissement de procédures de mise à jour et la gestion de l'historique des données administratives et scientifiques, la résolution des doublons (dans le cas de l'appartenance d'un même chercheur à deux organismes différents), le choix et la mise en œuvre des moyens de diffusion (logiciel d'interrogation et serveur videotex).
Enfin, l'une des tâches essentielles du comité a consisté à définir le contenu et la structure de la base.

L'INIST a mené l'étude comparative de la structure des fichiers internes des différents organismes. Débutée fin février 94, cette étude a révélé une grande disparité des bases internes de compétences, individuelles ou collectives, très différentes quant à leurs philosophies et leurs formats.

La maquette de la structure de la future base élaborée par le comité de Pilotage a été finalisée en juin 94. Elle se découpe en champs renseignables regroupés en blocs selon le modèle suivant :

- Identification de l'expert/coordonnées professionnelles.
- Activité et expertises :
  - Expertise antérieure (diplômes, qualifications validées, emplois antérieurs).
  - Activités actuelles (recherches en cours ou activités principales).
  - Définition de l'expertise (domaines scientifiques des recherches, champs de compétence, secteurs d'application industrielle économique ou de développement (dans le cas de l'ORSTOM), expertises diverses : brevets, activités d'enseignement, compétences linguistiques, expertises déjà effectuées, activités d'IST, appartenance à divers comités, etc.) (voir annexe 3).

Certaines zones sont renseignées par du texte libre et/ou des mots-clés. Les champs les plus significatifs sont bilingues français-anglais.

Il s'agissait, à partir d'un échantillon pertinent de curriculum vitae de construire une base de données globale et de la rendre, à titre expérimental, accessible pour évaluation.

Les fonctionnalités de cette base ont été évaluées par le comité de pilotage de FINES.

Les travaux se sont déroulés en trois grandes étapes :

1 - Préparation du fichier global.

2 - Chargement et mise en accès de cette base sur un serveur.

3 - Evaluation de la maquette.

II.1. PREPARATION DU FICHIER GLOBAL.

Recueil des informations :

Le CEA, le CNRS et l'INSERM disposent déjà de fichiers électroniques. Ils ont donc procédé à une extraction du nombre de curriculum vitae retenu, ce nombre devant être proportionnel au nombre total des chercheurs de chaque organisme.

L'INRA, et l'ORSTOM ne disposaient pas de fichiers déjà constitués. Ils ont fourni les renseignements correspondants aux C.V. selon le format qui leur a été indiqué par l'INIST après saisie des questionnaires sur papier ou sur bordereaux électroniques.

Le CIRAD n'ayant pas participé à l'étude préalable a fait l'objet d'une analyse détaillée de ses ressources et de la manière d'intégrer ses données.

Définition d'un format d'échange :

L'INIST a tout d'abord procédé à une étude approfondie de la structure des données dans les trois fichiers existants. Cette étape a comporté des réunions de travail bilatérales INIST et chacun des organismes concernés.

A partir de cette analyse, l'INIST a défini un format d'échange commun prenant en compte les caractéristiques de chacun des fichiers et basé sur le suivi des normes existantes.

Ce format s'appuie sur une structure balisée de type SGML (cf. annexe 1).

L'INIST a défini, pour les données devant être saisies par l'INRA, l'INRIA et l'ORSTOM, un cadre de saisie de ces informations.

Validation du format d'échange :

L'INIST a réuni les partenaires du projet afin de valider ses propositions et a procédé à des modifications en fonction des observations recueillies. La validation du contenu et de la structure des informations de la base a eu lieu en juin 95.

Ecriture des programmes de contrôle automatique des données :

L'INIST a mis en place des contrôles de validation automatiques reposant sur une DTD destinés à vérifier la conformité des fichiers reçus par rapport à la structure demandée : présence des champs obligatoires, longueur des champs, type de données : alphanumériques, etc.

Le développement, par la Direction Informatique de l'INIST, des programmes correspondant aux opérations décrites précédemment, c'est-à-dire l'écriture informatique de la DTD correspondant au balisage SGML, a été finalisé en juin 95. Il a permis de communiquer aux partenaires un outil de contrôle efficace.

II.2. TESTS, CHARGEMENT ET MISE EN ŒUVRE DE LA MAQUETTE.

Le choix et l'implémentation du système de gestion et d'interrogation de la maquette, selon une architecture client - serveur, ont été confiés par l'INIST au Bureau van Dijk (Ingénieurs Conseil en Gestion).

Jeux de tests :

Contrôle de la conformité des données des jeux de tests envoyés par les organismes.
Mise en conformité de ces données de jeux de tests et retour d'observations aux

organismes. Chacun des partenaires ayant toute liberté pour renseigner les différents champs en langage libre, il est apparu indispensable de procéder à un contrôle de la cohérence intellectuelle afin de réduire, autant que possible, les disparités sans modifier la nature des informations.
L'INIST a procédé à un repérage de ces disparités et proposé une solution pour en réduire l'importance.
Envoi des jeux de tests au Bureau van Dijk (septembre 95).
Retour d'observations aux organismes à la suite des tests réalisés par le Bureau van Dijk.

Mixage et validation des données :

Chacun des partenaires a fourni à l'INIST ses données dans le format défini, soit en procédant à une extraction à partir des fichiers existants, soit par saisie de questionnaires sur papier. L'INIST a pratiqué un contrôle de conformité, fin 95, sur ces notices avant de les transmettre au Bureau van Dijk.

Définition des modalités de recherche :

Compte tenu de l'origine diversifiée des fichiers et de la présence de plusieurs zones renseignées en langage libre, le logiciel retenu devait intégrer, pour une part importante, des possibilités d'interrogation en langage naturel.

Le Bureau van Dijk et l'INIST ont mené des études comparatives sur différents logiciels de ce type disponibles sur le marché ; le résultat de ces études a été analysé pour FINES.

Il a été également procédé à une analyse des systèmes d'interrogation de bases de type annuaire actuellement consultables, afin d'en mesurer les points forts et les points faibles.

Par ailleurs, l'INIST a recueilli les suggestions des partenaires du projet.

A partir des renseignements provenant de ces trois sources, il a été possible de passer à l'étape suivante.

Rédaction du cahier des charges du système de recherche documentaire:

Cette étape de définition des modalités et fonctions de la Recherche a consisté à détailler les spécifications attendues du SGBD nécessaire : champs interrogeables, index, préfiguration des écrans, etc.

Le projet a été présenté pour validation au Comité de Pilotage.

Choix et achat par le Bureau van Dijk d'une licence d'exploitation d'un SGBD relationnel :

Le choix du Bureau van Dijk et de l'INIST avait tout d'abord porté sur le système LEXIC de la Société GEDIMAGE, un produit séduisant de par ses caractéristiques très innovantes. Toutefois, suite au départ de son concepteur, il s'est avéré impossible de maintenir de manière fiable ce logiciel, et la décision a dû être prise fin 95 de se tourner vers un autre produit. Le logiciel AIRS de la Société EURITIS (European Information Technology for Information Science) a été retenu, parce que facile et rapide à paramétrer et mettre en œuvre et plus souple à modifier et maintenir (donc plus fiable pour une maquette) que d'autres produits envisageables (tel BASIS +).

AIRS, gestionnaire de bases de données riche en fonctionnalités, permet la recherche en texte intégral, l'utilisation de lexique, la création d'un thésaurus ; l'extraction est paramétrable, le stockage en SGML possible. De multiples restrictions d'accès sont utilisables. La recherche se fait par formulaire simple ou en mode multicritère, avec gestion de proximité à l'interrogation. L'aspect "historique" est bien développé. Ce logiciel est déjà utilisé pour la gestion documentaire du Ministère de l'Industrie, à l'INRS, à l'ADEME, à l'UNESCO et dans des centrales nucléaires entre autres.

Ce choix a été validé par le Comité de Pilotage en février 96.

Le Bureau van Dijk a effectué l'achat des licences d'exploitation.

L'autorisation de constitution de la maquette de cette base de données nominatives a été demandée à la CNIL et obtenue (arrêté du 25 octobre 95 en annexe 8).

Implantation de la maquette :

Le Bureau van Dijk a effectué, avec l'aide d'EURITIS, le paramétrage en fonction du cahier des charges précité et du format des notices. Ont été finalisés le déroulement des écrans ainsi que les conditions d'accès des participants à la maquette et la définition du protocole d'installation de celle-ci chez les membres du Comité de Pilotage (annexe 3).

Après réalisation des développements spécifiques de l'interface serveur, la phase des tests de mise en exploitation a été abordée.

689 notices ont été chargées en février 96 (annexe 5) ; l'insertion de ces données a fait apparaître certains problèmes d'écriture dont la synthèse a été communiquée aux organismes en mars 96 (annexe 2).

Les installations en local de la maquette pilotée par AIRS ont été effectuées en mars et début avril 96 chez chacun des organismes participant au projet.

L'application fonctionne selon un mode "pseudo client - serveur" sur un seul PC équipé de Windows 95.

Evaluation des résultats :

Le Bureau van Dijk a fourni à chaque participant un manuel d'utilisation de la maquette ainsi qu'un questionnaire (grille d'évaluation), élaboré avec l'INIST, permettant de recueillir les suggestions et critiques des évaluateurs (annexe 4).

L'évaluation s'est déroulée sur mars - début avril 96, sans thésaurus et avec verrouillage de l'administration des données. Elle porte surtout sur la structuration de la base et la cohérence sémantique et intellectuelle des données provenant des différents organismes.

La maquette a été présentée par l'INIST sur le stand du MENESR à IDT 96 (annexe 6).

Un document de synthèse, rédigé par l'INIST, faisant le point sur l'évaluation de la maquette suit le retour des observations des membres du Comité de Pilotage et sera soumis à celui-ci.

Annexe 1

Format de la base

DTD

III. SYNTHESE DES EVALUATIONS.

Se basant sur les retours d'évaluation de la maquette des membres du Comité de Pilotage, l'INIST a rédigé un rapport de synthèse qui servira d'outil pour les améliorations.

Dysfonctionnements du logiciel.

• Lignes illisibles
A l'affichage de zones de texte, certaines lignes sont illisibles.

• Nombre de notices
Apparemment, le nombre de CV présents dans la maquette ne correspond pas à ce qu'ont fourni les organismes. Ainsi :
– CIRAD : 101 notices fournies / 89 notices retrouvées
– INRIA : 35 notices fournies / 31 notices retrouvées
– INSERM : 201 notices fournies / 197 notices retrouvées
– INRA : 48 notices fournies / 44 notices retrouvées.

• Utilisation des index
Le logiciel "se plante" lorsque l'on interroge à l'aide d'un terme de longueur trop importante extrait du lexique.

• Dysfonctionnement de la recherche multi-critère
Nous aurions souhaité pouvoir entrer directement l'équation de recherche dans la zone "requête" située au bas de la boîte de dialogue de recherche multi-critère, soit pour modifier une équation, soit pour en saisir une nouvelle. Il nous semble que cette façon de formuler une question correspond au mode "expert" présenté dans la documentation. Malheureusement, nous n'avons pu utiliser ce mode comme le montre l'exemple ci-dessous :

étape 1 :
- ouverture de la boîte de dialogue de recherche multi-critère
- choix du champ act_def
- saisie de la valeur "BIOLOGIE" (sans les guillemets) dans la zone suivante située derrière le signe "–"
- lancement de la recherche (bouton "RECHERCHER")
- la zone "requête" affiche l'équation de recherche pendant l'exécution de la recherche.
  Cette équation est : ACT_DEF-"BIOLOGIE"
- la fenêtre s'ouvre et indique que 4 documents correspondant à cette équation.
  Jusqu'ici, le fonctionnement est normal.

étape 2 :
- la boîte de dialogue est restée ouverte. L'équation de recherche a disparu
- saisie manuelle de l'équation de la même recherche dans la zone "requête" : ACT_DEF-"BIOLOGIE"
- lancement de la recherche (bouton "RECHERCHER")
- un message d'erreur apparaît alors : "DFRsearchUA[Ret-1, ErrorVal-0]"
- ensuite le logiciel envoie le même message pour chaque recherche demandée quel que soit le mode opératoire. Il faut alors fermer le logiciel (plusieurs messages d'erreur) et le relancer.

Une fois sur deux, lorsque nous avons essayé de modifier une équation de recherche direment dans la zone "requête", le logiciel a eu ce comportement anormal. Après examen, il s'avère que ce problème peut être dû à de minimes erreurs de saisie.

Ergonomie

• Le logiciel manque encore d'habillage et de convivialité.
Ce fait s'explique par la rapidité avec laquelle il a été paramétré et implémenté, du fait du faible délai imparti. Cet inconvénient devrait être assez facilement levé.
Dans l'état, le logiciel n'est pas adapté – surtout pour la recherche multicritère – à une interrogation par des non-spécialistes. Les options de la boîte de dialogue de recherche multicritère sont difficilement compréhensibles. Le nombre d'opérations nécessaires à la construction d'une requête en utilisant les lexiques est trop important. L'accès au lexique par menu est trop lourd. Dans la recherche multicritère, il est trop facile de se tromper en saisissant un critère de recherche dans la zone de la requête au lieu de la fenêtre de saisie.
En résumé, la recherche multicritère risque de décourager les non-documentalistes.

• Visualisation par effet vidéo inverse.
Il serait intéressant d'en bénéficier pour les requêtes avec des termes tronqués. De même, lors de la visualisation des fiches de réponse par champs, les termes de requête n'apparaissent pas en brillance inversée si l'accentuation ou la police sont différentes entre la requête et la réponse.

• Il est indispensable de verbaliser les noms des champs en recherche multicritère.

• Les résultats "par tableaux" sont difficiles à lire.

• Prévoir un bouton "arrêt de la recherche en cours".

• Il y a trop de fenêtres, en particulier en multicritère. La gestion des fenêtres de requête est très incommode ; celles-ci cachent les commandes de visualisation des références trouvées et doivent être fermées après chaque recherche.

• En page principale de la requête par formulaire, le mode d'emploi des ET/OU et des troncatures devrait être explicité ou accessible facilement depuis cette page.

• La présence de l'historique est très utile.

Fonctionnalités de recherche

• Il est souhaitable de créer un basic index reprenant les rubriques à contenu scientifique (activité, secteurs d'application, publications, enseignement, brevets, etc.). Ce basic index pourrait être utilisé dans une rubrique "compétences" unique du mode de requête par formulaire, celle-ci ne s'appuyant pas actuellement sur les champs les plus informatifs. La rubrique "expériences antérieures" est à inclure dans cette rubrique.

• Dans la requête par formulaire, la troncature droite devrait être implicite.

• Requête par formulaire.
L'exemple simple décrit ci-après à l'aide de la commande "Recherche par formulaire" illustre un risque d'erreur banal mais qui peut être source de silence dans les réponses.
- Requête ou étape n° 1 : ***MATERIAUX,COMPOSITES*** donnent cinq réponses.
- Requête ou étape n° 2 : ***MATERIAUX, COMPOSITES*** (avec un caractère blanc derrière la virgule) donnent trois réponses.
En visualisant l'historique, vous pouvez modifier directement la requête ou l'étape de votre choix.

On peut lire pour l'étape n° 1 la description suivante :

(ACT_DEF=MATERIAUX ET ACT_DEF=COMPOSITES)
OU (ACT_IND=MATERIAUX ET ACT_IND=COMPOSITES).

On peut lire pour l'étape n° 2 la description suivante :

((ACT_DEF=MATERIAUX) ET (ACT_DEF=COMPOSITES))
OU (ACT_IND=MATERIAUX ET ACT_IND=COMPOSITES).

A la lecture de la formule de l'étape n° 2, la position des parenthèses supplémentaires ne me semble pas devoir conduire à des résultats différents de ceux de l'étape n° 1. Cependant, il y a bien un écart sur le nombre de réponses. Si je modifie directement l'étape n° 1 pour la rendre identique à la formule de l'étape n° 2, on crée donc une étape n° 3 qui donne les résultats de l'étape n° 1 et non ceux de l'étape n° 2 comme on pourrait s'y attendre. On fait le même constat si on modifie directement l'étape n° 2 pour la rendre identique à la formule de l'étape n° 1, à savoir une étape n° 4 avec les résultats de l'étape n° 2 et non ceux de l'étape n° 1.

En conclusion, même si ça simplifie la stratégie de recherche (en temps de reformulation voisine en principe, et surtout si la requête est complexe), il n'est pas prudent de réutiliser une étape de recherche pour la modifier, quand on n'est pas totalement sûr de son origine et de ses résultats.

• Il est dommage que l'on ne puisse pas copier simultanément plusieurs lignes du lexique directement dans l'équation de recherche.

Implémentation

Un certain nombre de critiques relèvent de l'implémentation des données ; des propositions sont faites :

- éclater les index "publications" selon :
  - type
  - titre
  - année

A noter que l'on peut individualiser ces champs années et les rendre numériques, mais l'on perd alors le lien avec les publications, sauf à l'affichage.

- éclater les index "emplois antérieurs" :
  - description fonction
  - géographie.

Même remarque en ce qui concerne les coopérations pour faire apparaître une liste géographique ; ces deux listes pouvant éventuellement être fusionnées.

- éclater le champ "langues" de façon à faire apparaître une liste de référence des langues.
- éclater le champ "identité" de façon à faire apparaître une liste de noms.
- De manière générale, il semble préférable d'éclater les mots des lexiques à contenu scientifiquement pertinent en unitermes, tout en conservant les pluritermes,

ex. "infarctus du myocarde" —> "infarctus"
"myocarde"
"infarctus du myocarde"

par contre, une requête sur "infarctus du myocarde" ne devrait pas générer des réponses avec "infarctus" individualisé par effet vidéo inverse, même si les deux termes sont présents dans le CV.

Est-il possible, enfin, d'activer des liens hypertextes sur les organismes, les mots-clés et les coauteurs des publications afin de naviguer d'un CV à l'autre sans reformuler de requête ?

Cette hypothèse n'apparaît pas envisageable dans la configuration actuelle du logiciel.

Qualité de la saisie des données.

- Hétérogénéité gênante des polices de caractères d'un CV à l'autre.
- Erreurs typographiques assez fréquentes, ce qui génère du silence si l'on n'utilise pas les index.
- L'accolement de mots dans les zones de texte libre pose aussi problème à l'interrogation. Il pourrait s'agir d'un problème de retour chariot mal interprété (parseur ?).
- L'absence de blancs entre les mots-clés se traduit aussi par du silence.

Gestion du vocabulaire

- L'absence de hiérarchisation du vocabulaire est une cause assez importante de silence. Le thésaurus serait une solution, mais il apparaît plus intéressant de réfléchir à une interface intelligente. Ce type d'outil permettrait de nettoyer les index, ce qui s'avère indispensable même sur un volume aussi faible de CV.

- La création d'un basic index où les pluritermes seront éclatés en unitermes semble indispensable. Malheureusement, il ne semble pas possible d'obtenir directement de basic index à l'affichage.

- L'indexation est dans son ensemble de bonne qualité mais inégale et peu cohérente d'un organisme à l'autre, si ce n'est pas dans le même organisme.

- L'opportunité de créer un plan de classement a été évoquée, ceci afin de pouvoir restreindre les sessions d'interrogation à un domaine défini. Cette démarche est cependant réductrice par rapport à une des applications envisageables de la base qui est de permettre de retrouver l'expertise là où on ne s'attend pas forcément à la trouver. Une structuration interne de la base en domaines interdirait aussi d'effectuer des recherches croisées (polythématiques), d'intérêt certain.
Beaucoup plus valorisant serait l'ajout d'un outil de cartage de l'information. C'est le cas, par exemple, de Notizer DLL qui travaille sur texte intégral pour l'étude de pertinence et la recherche de concepts.

- L'utilisation de symboles d'éléments en dehors des formules chimiques est à bannir.

Niveau de structuration de la base et contenu des champs.

La structuration apparaît globalement suffisante, une restructuration plus fine n'est pas nécessaire, à condition de constituer des index individualisés sur les champs les plus signifiants de certaines rubriques. De toute façon, on n'envisage pas de remettre en cause la structuration SGML de FINES, mais plutôt la modélisation de la base AIRS.

A l'exception de la rubrique "Secteurs d'application", la cohérence du contenu des rubriques apparaît à l'ensemble des organismes globalement satisfaisante. Le taux de réponse est bon et il serait dommage – car appauvrissant – de chercher à homogénéiser de trop les CV.

- Le plus gros problème est posé par les secteurs d'application de l'expertise : les règles définissant les contenus de ce champ restent entièrement à redéfinir. Les mots-clés doivent-ils être génériques ou spécifiques ? Un texte libre est-il nécessaire ? L'ambiguïté est totale entre les mots-clés de DOMFRA et ceux d'INDEXFRA.

- Rubrique "Identité de l'expert" : le remplissage est hétérogène d'un organisme à l'autre. L'énoncé des titres, en particulier, relève parfois des distinctions scientifiques ou des diplômes. Il ne doit s'agir que d'un élément de l'adresse : "Docteur", "Monsieur le professeur".

- Rubrique "adresse professionnelle" : tout comme dans le cas de l'adresse de la personne à contacter, il faut une adresse complète.

A noter que ces rubriques d'identité et d'adresse de l'expert devraient peut-être être déclarées facultatives si l'on souhaite que l'expert ne puisse être approché directement ; ce problème pourrait être résolu par les fonctionnalités de restriction d'accès qu'offre AIRS.

- Il est indispensable de verbaliser les codes ISO des noms de pays, ainsi que de rendre homogène l'écriture de ces noms dans les rubriques "Coopérations" et "Emplois antérieurs" : pour un même organisme, on trouve tantôt le développé complet du nom, tantôt un code ISO.

- L'unité de durée doit être mentionnée pour les rubriques "coopérations" et "enseignement dispensé".

Comparaison des principales fonctions des logiciels AIRS et FOLIO

FONCTION	AIRS	FOLIO
Sécurité et confidentialité / accès par mot de passe	oui	possible
Type de recherche		
- thesaurus	possible	non
- texte intégral	oui	oui
- index	oui	oui
- séparateurs	oui	oui
- recherche multicritères	oui	oui
- recherche multichamps	oui	oui
- recherche multibases	oui	oui
- recherche par synonymes	possible	oui
- recherche par racines	possible	oui
- historique	oui	oui
Opérateurs		
- masques	oui	oui
- troncature gauche	oui	oui
- troncature droite	oui	oui
- parenthèses	oui	oui
- opérateurs booléens	oui	oui
- opérateurs de proximité	oui	oui
- opérateurs mathématiques	oui	oui
Editions		
- du document à l'écran	oui	oui
- de l'ensemble des documents	oui	oui
- des documents marqués	non	oui
- export sous forme de fichier	oui	oui

Maquette FINES : spécifications techniques

Introduction

Le présent document développe et explicite les spécifications que devront suivre les fichiers que nous fourniront les organismes participant au projet FINES, fichiers reformatés par leurs soins conformément à ces spécifications.

Ces fichiers, une fois concaténés, constitueront, avec l'outil de gestion et d'interrogation approprié, la future maquette FINES.

Support physique et format des données

Le support préférentiel de transfert de l'information sera la disquette $3^{1/2}$".
Les fichiers, de type séquentiel, utiliseront le mode ASCII étendu.

Structuration des champs

Les éléments d'information, définis dans les spécifications de format, sont structurés à l'aide d'un balisage de type SGML.

La structure est explicitée au moyen d'une arborescence ci-jointe.

Le début de chaque unité d'information sera balisé par le nom de l'élément correspondant, entre < > ; et les fins d'unités par le nom de l'élément, entre < / >.

Dans le cas de multioccurrences, il conviendra de répéter le balisage de début et de fin de l'élément concerné à chaque occurrence. Ainsi, par exemple :

> <expdef> <domaine> ferrite </domaine> <domaine> grenat </domaine>
> <domaine> magnétisme </domaine> <domaine> couches minces </domaine>
> <domaine> pulvérisation cathodique </domaine> </expdef>

Les éléments faisant l'objet d'une syntaxe attribut devront obéir à la règle d'écriture suivante :

> <langue ^ att = "lu" ^ att = "1">

Les champs déclarés obligatoires seront systématiquement alimentés, la notice faisant, dans le cas contraire, l'objet d'un rejet et d'une procédure de correction.

Lorsqu'une information n'est pas fournie, les balises concernant l'élément correspondant n'apparaissent pas.

A l'intérieur des textes libres, les paragraphes peuvent être matérialisés par des balises de début et de fin : <P> et </P>. Ce balisage n'est pas obligatoire.

Exemple de balisage du bloc "identité"

<identite> <nom> DUPONT </nom> <titreid> Dr </titreid>
<prenom> Jean-Paul </prenom> <fonction> DR2 </fonction> </identite>

Exemple de balisage d'un bloc "identité" incomplet

<identite> <nom> DUPONT </nom> <prenom> Jean-Paul </prenom>
<fonction> DR2 </fonction> </identite>

La balise <titreid> n'est pas mentionnée, car non porteuse d'information.

Chaque curriculum vitæ est délimité par une balise de type <record>.

FINES / SPECIFICATIONS DU FORMAT D'ÉCHANGE FINES

Eléments de données	Balises	Attributs	Syntaxe	Oblig./ facult.	Répéti -tivité	Longueur	Notes et Recommandations
Identité de l'expert	IDENTITE		-	0	-	-	Pour des raisons de confidentialité, ce bloc identité peut être masqué (par des caractères "x") ; se reporter alors au bloc "personne à contacter".
Titre civil	CIVIL		M. ou MME.	0	-	4	Majuscules, 2 ou 4 caractères.
Nom	NOM		Libre	0	-	variable	Patronyme de l'expert en majuscules.
Prénoms	PRENOM		Prénom(s) en entier ou initiales de prénom	0	-	variable	Prénom usuel de l'expert ; les prénoms multiples sont possibles. Prénoms ou initiale(s) en majuscules.
Titre (s) universitaires, professionnels	TITREID		Libre	F	-	variable	Titre de préférence sous forme développée. Ex : Docteur. Professeur.
Fonction ou grade	FONCTION		Libre	F	-	variable	Il s'agit soit du poste actuellement occupé (ex : chargé de recherche 1ère classe CNRS ; chef de service), soit du grade, de préférence en clair (ex : Directeur de recherche 2)..
Adresse professionnelle	ADRESSE			0	-	-	
Nom de l'Etablissement	NOMETAB		Nom de l'Etablissement en clair	0	-	variable	Nom sous forme développée de l'Etablissement rattaché à l'organisme (ex : Centre d'Etudes de Saclay).
N°, rue, Boîte Postale	RUE		Libre	0	-	Variable	Numéro, nom de la rue, boîte postale. Chaque élément à séparer par des virgules.
Code postal	CODPOST		- 5 chiffres (France) - Libre (étranger)	0	-	Variable	
Ville	VILLE		Ville	0	-	Variable	De préférence en majuscules. Ex : PARIS
Cedex ou secteur postal	CEDEX		Cedex ou secteur postal	F	-	Variable	Ex : CEDEX 06 ou 07 SP
Pays	PAYS		Code ISO 3166	0	-	3	Ex : FRA pour France (majuscules).
N° de téléphone	TEL		Code pays + N° national	F	-	16	Format international. Ex : +31-43-304050, +33-1-69823153 (aucun "blanc", signes autorisés : - et +).
N° de fax	FAX		idem TEL	F	-	16	
Adresse courrier électronique	EMAIL			F	-	Variable	Intitulé complet de l'adresse électronique, avec @ ou (at).

Eléments de données	Balises	Attributs	Syntaxe	Oblig./ facult.	Répéti -tivité	Longueur	Notes et Recommandations
Organisme duquel dépend l'expert	ORGAN		-	0	-	-	
Nom de l'organisme	NOMORGAN		Nom sous forme développée	0	-	Variable	Nom, sous forme développée, de l'organisme dont dépend l'établissement. (Ex : Commissariat à l'Energie Atomique).
Sigle de l'organisme	SIGLE		Sigle sans point	0	-	Variable	En majuscules et sans point. ex : CEA
Code du laboratoire ou de l'unité de recherche	CODE			F	-	Variable	ex : URA 7248 ; U306 Imbrication de codes possible précisant la hiérarchie administrative des services (cf. CEA).
Intitulé de l'unité de recherche ou du laboratoire	UNIRECH		Libre	0	-	Variable	Ex : - UR agronomie - Centre de Recherche sur l'Habitat - Génie Biomédical et Diabète Sucré.
Champs actuels d'expertise ou de compétence	ACTIVITE		-	0	-	-	
Définition (en français)	DEFFRA		Libre	0	-	Variable	Texte libre en français décrivant les intérêts thématiques actuels : champs d'expertise ou de compétence, activités scientifiques et techniques actuelles, thèmes de recherches en cours. Jusqu'à 3000 caractères.
Définition (en anglais)	DEFENG		Libre	F	-	Variable	Idem DEFRA, mais en Anglais.
Indexation (en français)	INDEXFRA		Mots-clés	F	30	Variable	Mots-clés français décrivant les intérêts thématiques et les activités scientifiques et techniques actuels. Jusqu'à 30 mots-clés.
Indexation (en anglais)	INDEXENG		Mots-clés	F	30	Variable	Idem INDEXFR, mais en anglais.
Autre(s) expertise(s)	AUTEXP			F	-		
Autre(s) expertise(s) (en français)	AUTEXPFR		Libre	0	-	Variable	Texte libre en français décrivant d'autres aspects de l'expertise hors activité principale. Jusqu'à 1500 caractères.
Autre(s) expertise(s) (en anglais)	AUTEXPEN		Libre	F	-	Variable	Idem AUTEXPFR, mais en anglais.

Eléments de données	Balises	Attributs	Syntaxe	Oblig./ facult.	Répéti -tivité	Longueur	Notes et Recommandations
Secteur	SECTEUR		-	0	3		Secteur(s) d'application des compétences.
Domaine (en français)	DOMFRA		Mots-clés ou texte libre	F	-	Variable	Texte libre (ou mot-clé générique) en français définissant les domaines scientifiques et techniques d'application des activités ou des recherches. Jusqu'à 1500 caractères. Sans abréviation.
Domaine (en anglais)	DOMENG		Mots-clés ou texte libre	F	-	Variable	Idem DOMFRA, mais en anglais.
Indexation (en français)	INDEXFRA		Mots-clés	F	20	Variable	Mots-clés en français définissant les secteurs possibles d'application des recherches (industrie, médecine, développement). Jusqu'à 20 mots-clés.
Indexation (en anglais)	INDEXENG		Mots-clés	F	20	Variable	Idem INDEXFRA, mais en anglais.
Diplôme(s)	DIPLOME		-	F	3	-	Le diplôme de niveau le plus élevé dans chaque domaine de compétences 3 diplômes différents sont possibles.
Intitulé du diplôme	INTITULE		Intitulé en clair du diplôme	0	-	Variable	Intitulé officiel du diplôme, sous forme développée, tel que donné par l'expert ou verbalisation d'un code (ex : V = doctorat en médecine vétérinaire).. Une liste des intitulés officiels sera fournie.
Lieu d'obtention du diplôme	LIEU		Libre	F	-	Variable	Nom de l'Université, de l'Ecole ou, à défaut, de la ville ou le diplôme a été obtenu (cf. liste ci-jointe). Une liste des noms d'Université ou d'Ecoles sera fournie.
Année d'obtention du diplôme	DATE		Code ISO 8601	F	-	4	Indication de l'année.
Activités antérieures	EMPLOI		-	F	5	-	Les 5 dernières activités exercées, au sein d'organismes différents, ou bien au sein du même organisme, si l'activité a évolué sensiblement dans sa définition ou s'est traduite par divers détachements à l'étranger (pour des activités d'aide au développement, par exemple).
Période	PERIODE		Année ou fourchette (ex : 1984-1988)	0	-	Variable	Durée de chaque emploi (ex : 1992-1994).
Organisme	NOMORGAN		Nom sous forme développée	0	-	Variable	Nom, sous forme développée, de l'organisme dans lequel l'emploi a été exercé.
Lieu	LIEU		Libre	F	-	Variable	Ville, région ou état.

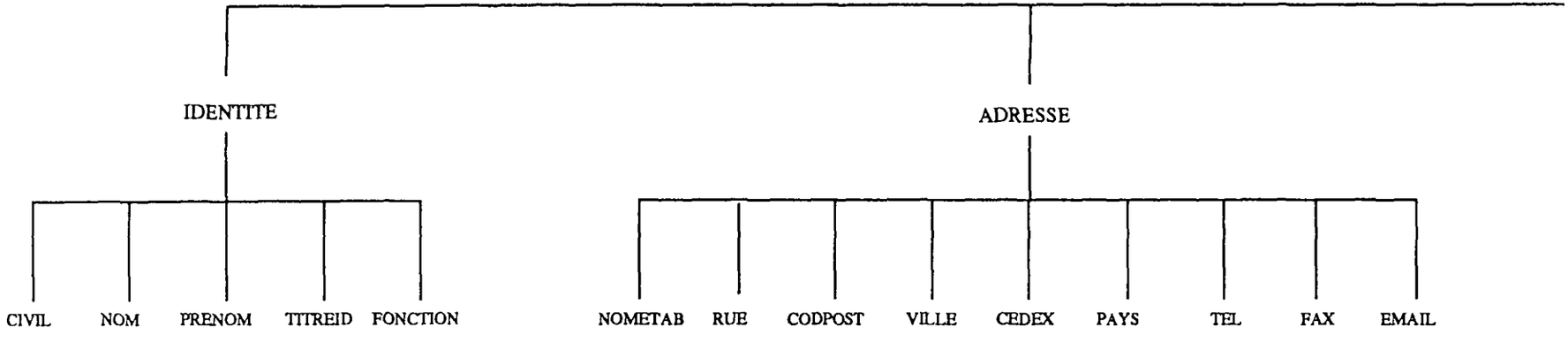
Eléments de données	Balises	Attributs	Syntaxe	Oblig./ facult.	Répéti -tivité	Longueur	Notes et Recommandations
Pays	PAYS		Code ISO 3166	0	-	3	Pays ou l'emploi a été exercé. Utiliser le code ISO (ex : ITA pour Italie).
Qualité	QUALITE		Libre	F	-	Variable	Qualité de l'expert pendant la durée de l'emploi (ex : ingénieur d'étude).
Description de l'activité (en français)	DESFRA		Libre	F	-	Variable	Description thématique de l'activité exercée, en français. Jusqu'à 1000 caractères.
Description de l'activité (en anglais)	DESENG		Libre	F	-	Variable	Idem DESFRA, mais en anglais.
Publications	PUBLI		-	F	12	-	Les plus significatives et les plus récentes dans le cadre de la définition de l'expertise.
Type de publications	TYPE		Code	0	-	1	"L" pour ouvrage, "A" pour article, "C" pour communication de congrès, "R" pour rapport et "S" pour autres.
Titre original de l'objet de publication	TITRE		Libre	0	-	Variable	Titre original de la publication y compris les sous-titres, le cas échéant.
Langue du titre	LTITRE			F	-	3	Selon la norme. Ex. : FRE, ENG, GER, ITA, SPA, etc...
Titre traduit	ALTITRE		Libre	F	-	Variable	Titre traduit, le cas échéant.
Langue du titre traduit	LALTITRE			F	-	3	Selon la norme. Ex. : FRE, ENG, GER, ITA, SPA, etc...
Auteurs	COAUT		Nom et prénom(s) si possible, sinon nom et initiales.	F	-	Variable	Auteurs: nom et prénom(s) ou initiales ou nom d'une collaboration d'auteurs, en majuscules. Le séparateur est la virgule.
Date	DATE		ISO 8601	0	-	4	Date de parution
Autre(s) information(s)	INFPUB			F	-	Variable	Faire figurer dans cette zone les différents éléments (séparés par des virgules) sources de la publication : si "L" éditeur et lieu d'édition si "A" titre de la revue, ISSN, n° de volume, pagination si "C" nom, lieu du congrès, éditeur ou organisateur. si "R" n° de rapport, commanditaire
Brevet	BREVET		-	F	3	-	Les trois brevets les plus caractéristiques de l'expertise
Titre	TITRE		Libre	0	-	Variable	Le titre officiel du brevet
Co-inventeur(s)	INVENT		Noms et prénoms ou initiales	F	5	Variable	Co-inventeurs du brevet.

Eléments de données	Balises	Attributs	Syntaxe	Oblig./ facult.	Répéti -tivité	Longueur	Notes et Recommandations
N° de brevet/N° de demande	NOBREV		Syntaxe normalisée	0	-	15	
Année de dépôt	DATE		Code ISO 8601	0	-	4	
Pays couvert(s)	PAYSCOUV		Libre	F	-	Variable	Ex : Europe
Déposant(s) du brevet	DEPOSANT		Libre	F	-	Variable	Industrie ou institution
Distinctions scientifiques, Activités associatives ou officielles	DISTSCI		Texte libre	F	-	Variable	Appartenance à des corps professionnels, à des sociétés savantes, groupes interministériels, conseils ou commissions scientifiques, groupes de réflexion, conseils d'administration de sociétés, comités de rédaction ; expertises validées officiellement, prix et distinctions scientifiques., etc.
Coopérations, partenariats ou missions	COOPE		-	F	5	-	Participation à des coopérations nationales et internationales, missions de courte durée dans les domaines industriels, de la médecine ou du développement ; ou partenariat en cours.
Intitulé	INTITULE		Libre	0	-	Variable	Intitulé développé des objectifs.
Organisme concerné	NOMORGAN		Nom sous forme développée	0	-	Variable	Nom sous forme développée de l'organisme partenaire.
Lieu	LIEU		Libre	0	-	Variable	Ville, région ou état.
Pays	PAYS		Code ISO 3166	0	-	3	Ex ITA pour Italie.
Durée	DUREE		Libre	F	-	Variable	Durée prévue en mois.
Date	DATE		Code ISO 8601	F	-	4	Indication de l'année de début.
Enseignement	ENSEIGN		-	F	-	-	Activités de formation et d'enseignement les plus en rapport avec l'expertise.
Cours dispensé	COURS		-	F	3	-	
Intitulé ou description	*INTITULE*		Libre	0	-	Variable	Dénomination officielle des cours dispensés ou texte libre les décrivant.
Localisation	*LIEU*		Syntaxe libre	0	-	Variable	Nom de l'Université ou de l'Ecole (ex : Nancy I)
Durée	*DUREE*		Nombre d'heures	F	-	Variable	

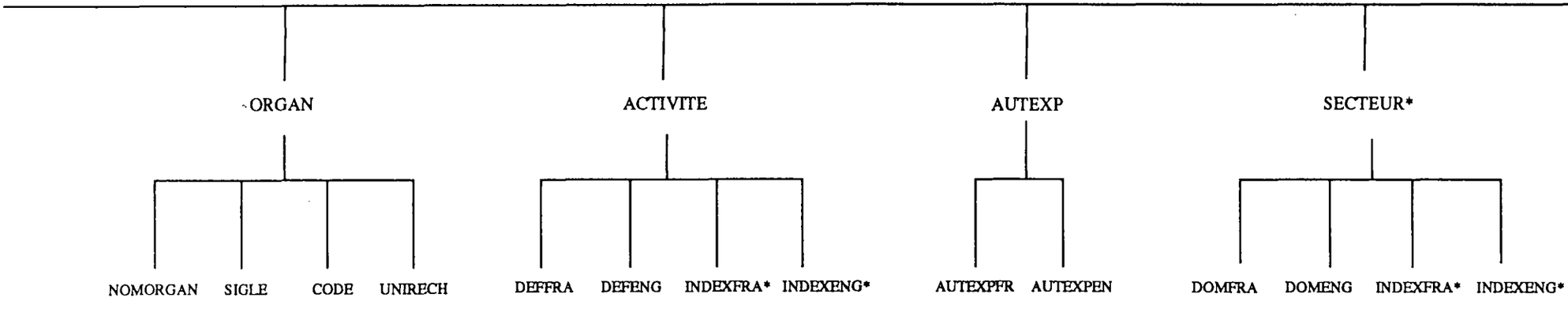
Eléments de données	Balises	Attributs	Syntaxe	Oblig./ facult.	Répéti -tivité	Longueur	Notes et Recommandations
Date	*DATE*		Syntaxe conseillée : "1990" ou "1991-1993" ou "depuis 1992"	F	-	Variable	
Encadrement de thèses	ENCADR		-	F	3	-	Les trois activités d'encadrement de thésard les plus en rapport avec la définition de l'expertise.
Titre de la thèse	*TITTHES*		Syntaxe libre	0	-	Variable	Titre officiel de la thèse.
Localisation	*LIEU*		Syntaxe libre	F	-	Variable	Nom de l'Université ou de l'Ecole (ex : Nancy I)
Année de soutenance	*DATESOUT*		Code ISO 8601	F	-	4	Année de soutenance (effectuée ou prévue).
Activités de communication (IST)	IST		Libre	F	-	Variable	Activités diverses d'IST : vulgarisation, exposition, interviews, documentation, participation à la constitution de bases de données de fichiers ou répertoire de données, réseaux documentaires.
Compétences linguistiques	LANGUES		-	0	6	-	Y compris langue maternelle.
Langue	LANGUE		Libre	0	-	Variable	Ex : Français
Lu	LU		0, 1, 2 ou 3	F	-	1	4 valeurs possibles de 0 à 3 : 0 (rudiments), 1 (connaissance de base), 2 (pratique moyenne), 3 (pratique courante).
Ecrit	ECRIT		0, 1, 2 ou 3	F	-	1	Id.
Parlé	PARLÉ		0, 1, 2 ou 3	F	-	1	Id
Personne à contacter	CONTACT			0	-	-	Responsable à contacter pour toute demande d'information complémentaire (l'expert lui-même ou une autre personne).
Titre civil de la personne à contacter	CIVIL		M. ou MME.	0	-	4	Majuscules ; de 2 à 4 caractères.
Nom de la personne à contacter	NOM		Libre	0	-	Variable	En majuscules.
Prénoms de la personne à contacter	PRENOM		Prénom(s) en entier ou initiales de prénom	0	-	Variable	Prénom usuel de la personne à contacter ; les prénoms multiples sont possibles. Prénoms ou initiale(s) en majuscules.
Titre(s) universitaire(s) professionnel(s) de la personne à contacter	TITREID		Libre	F	-	Variable	Titre, de préférence sous forme développée. Ex. : Docteur, Professeur.

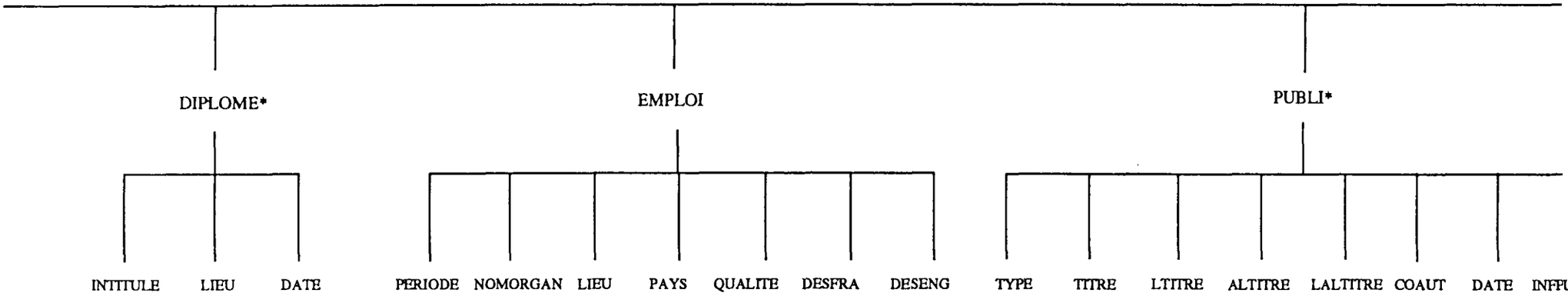
Eléments de données	Balises	Attributs	Syntaxe	Oblig./ facult.	Répéti -tivité	Longueur	Notes et Recommandations
Titre ou fonction de la personne à contacter	FONCTION		Libre	F	-	Variable	Titre ou fonction sous forme développée, de préférence.
Nom de l'Etablissement	NOMETAB		Nom de l'Etablissement en clair	0	-	variable	Nom sous forme développée de l'Etablissement rattaché à l'organisme (ex : Centre d'Etudes de Saclay).
N°, rue, Boîte Postale	RUE		Libre	0	-	Variable	Numéro, nom de la rue, boîte postale. Chaque élément à séparer par des virgules.
Code postal	CODPOST		- 5 chiffres (France) - Libre (étranger)	0	-	Variable	
Ville	VILLE		Ville	0	-	Variable	De préférence en majuscules. Ex : PARIS
Cedex ou secteur postal	CEDEX		Cedex ou secteur postal	F	-	Variable	Ex : CEDEX 06 ou 07 SP
Pays	PAYS		Code ISO 3166	0	-	3	Ex : FRA pour France (majuscules).
N° de téléphone	TEL		Code pays + N° national	F	-	16	Format international. Ex : +31-43-304050, +33-1-69823153 (aucun "blanc", signes autorisés : - et +).
N° de fax	FAX		idem TEL	F	-	16	
Adresse courrier électronique	EMAIL			F	-	Variable	Intitulé complet de l'adresse électronique, avec @ ou (at).
Gestion de la notice	GESTION			0	-	-	
Date de la première entrée FINES	DATEENTR		Code ISO 2014	0	-	8	
Date de la dernière mise à jour FINES	DATEMAJ		Code ISO 2014	0	-	8	
N° de référence de la notice FINES	NOREF		Numérique	0	-	5	Numéro séquentiel précédé du sigle de l'organisme.

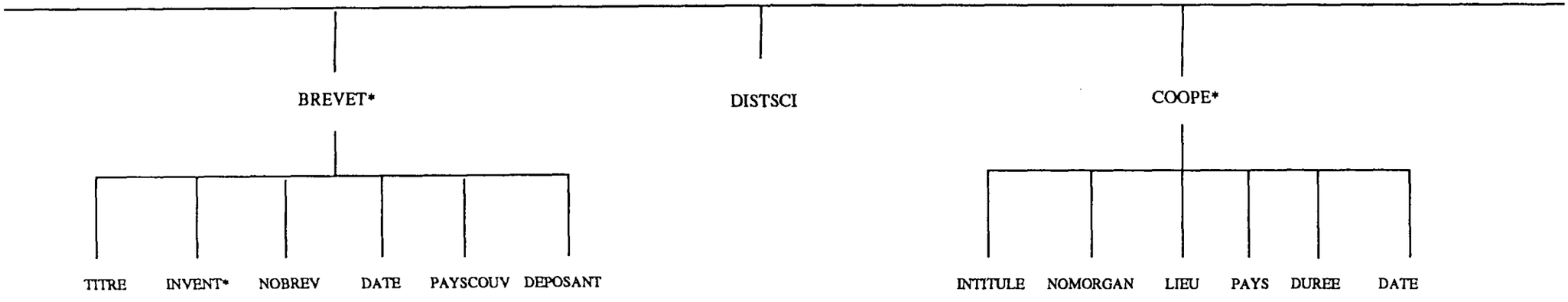
BASE FINES
ARBORESCENCE

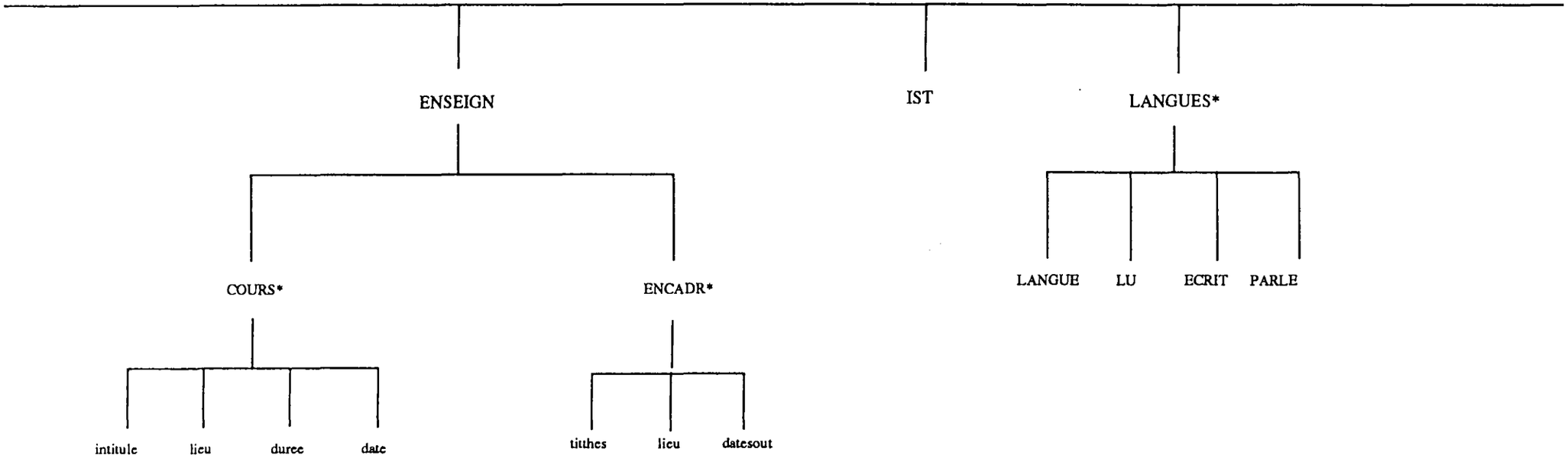


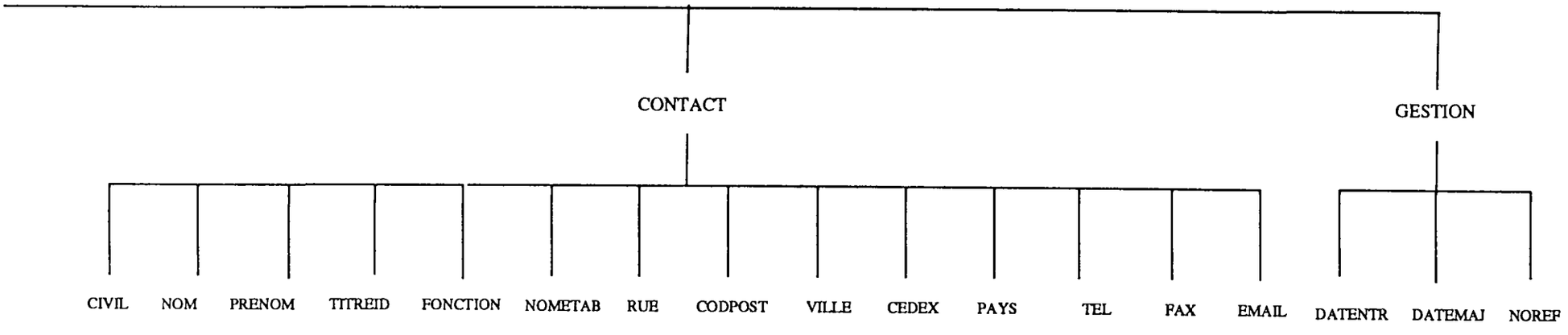
Note : L'astérisque (*) signale les zones répétitives.


ORGAN
NOMORGAN
SIGLE
CODE
UNIRECH
ACTIVITE
DEFFRA
DEFENG
INDEXFRA*
INDEXENG*
AUTEXP
AUTEXPFR
AUTEXPEN
SECTEUR*
DOMFRA
DOMENG
INDEXFRA*
INDEXENG*


DIPLOME*
EMPLOI
PUBLI*
INTITULE
LIEU
DATE
PERIODE
NOMORGAN
LIEU
PAYS
QUALITE
DESFRA
DESENG
TYPE
TITRE
LTITRE
ALTITRE
LALTITRE
COAUT
DATE
INFPUB


BREVET*
DISTSCI
COOPE*
TITRE
INVENT*
NOBREV
DATE
PAYSCOUV
DEPOSANT
INTITULE
NOMORGAN
LIEU
PAYS
DUREE
DATE


ENSEIGN
IST
LANGUES*
COURS*
ENCADR*
LANGUE
LU
ECRIT
PARLE
intitule
lieu
duree
date
titres
lieu
datesout


CONTACT
GESTION
CIVIL
NOM
PRENOM
TITREID
FONCTION
NOMETAB
RUE
CODPOST
VILLE
CEDEX
PAYS
TEL
FAX
EMAIL
DATENTR
DATEMAJ
NOREF

```
<!DOCTYPE FINES [
<!--
   DTD Fines Version 1.0
   Creation D. Lechaudel JUIN 95
   Copyright INIST/CNRS.
   Invocation :
   <!DOCTYPE FINES PUBLIC "-//INIST//DTD fines//FR" >
   Cette version fait suite a la version 0.1 de MAI 95
   les modifications sont liees aux changements de specifications
   decidees lors du comite de pilotage du 28/6/95
   - passage de tous les attributs en champs propres
   - creation des champs LTITRE, LALTITRE dans PUBLI
   - creation du champ TITREID dans CONTACT
   l'auteur degage toute responsabilite quant a la structuration non SGML des donnees
   contacter le comite de pilotage FINES en cas de probleme

-->

        <!-- ENTITIES   -->
<!ENTITY % doctype "FINES"   -- type de document Fines -->
<!ENTITY % un.txt   "#PCDATA"  -- unité textuelle -->
<!ENTITY % m.niv    "#PCDATA"  -- niveau de la langue -->

        <!-- ELEMENTS   -->
<!--    ELEMENTS     MIN  CONTENT -->
<!ELEMENT %doctype;     - - (RECORD+)  >

<!ELEMENT P           - O (%un.txt;)      -- Paragraphe -->
<!ELEMENT RECORD      - -
(IDENTITE,ADRESSE,ORGAN,ACTIVITE,AUTEXP?,SECTEUR+,DIPLOME*,EMPLOI*,

PUBLI*,BREVET*,DISTSCI?,COOPE*,ENSEIGN?,IST?,LANGUES+,CONTACT,GESTION)
+(P) >
<!ELEMENT IDENTITE    - - (CIVIL,NOM,PRENOM,TITREID?,FONCTION?)    -- Identite
de l'expert -- >
<!ELEMENT CIVIL        - - (%un.txt;)    -- Titre civil -- >
<!ELEMENT NOM          - - (%un.txt;)    -- Nom -- >
<!ELEMENT PRENOM       - - (%un.txt;)    -- Prenoms -- >
<!ELEMENT TITREID      - - (%un.txt;)  -- Titre (s) universitaires, professionnels -- >
<!ELEMENT FONCTION     - - (%un.txt;)  -- Fonction ou grade -- >

<!ELEMENT ADRESSE      - -
(NOMETAB,RUE,CODPOST,VILLE,CEDEX?,PAYS,TEL?,FAX?,EMAIL?)    -- Adresse
professionnelle -- >
<!ELEMENT NOMETAB      - - (%un.txt;)    -- Nom de l'Etablissement -- >
<!ELEMENT RUE          - - (%un.txt;)    -- Nø, rue, Boite Postale -- >
<!ELEMENT CODPOST      - - (%un.txt;)    -- Code postal -- >
<!ELEMENT VILLE        - - (%un.txt;)    -- Ville -- >
<!ELEMENT CEDEX        - - (%un.txt;)  -- Cedex ou secteur postal -- >
<!ELEMENT PAYS         - - (%un.txt;)    -- Pays -- >
<!ELEMENT TEL          - - (%un.txt;)  -- Nø de telephone -- >
<!ELEMENT FAX          - - (%un.txt;)  -- Nø de fax -- >
<!ELEMENT EMAIL        - - (%un.txt;)  -- Adresse courrier electronique -- >

<!ELEMENT ORGAN        - - (NOMORGAN,SIGLE,CODE?,UNIRECH)    -- Organisme
duquel depend l'expert -- >
<!ELEMENT NOMORGAN     - - (%un.txt;)    -- Nom de l'organisme -- >
<!ELEMENT SIGLE        - - (%un.txt;)    -- Sigle de l'organisme -- >
<!ELEMENT CODE         - - (%un.txt;)  -- Code du laboratoire ou de l'unite de recherche --
>
```

```
<!ELEMENT UNIRECH      - - (%un.txt;)    -- Intitule de l'unite de recherche ou du
laboratoire -- >

<!ELEMENT ACTIVITE     - - (DEFFRA,DEFENG?,INDEXFRA*,INDEXENG*)    -- Champs
actuels d'expertise ou de competence -- >
<!ELEMENT DEFFRA       - - (%un.txt;)    -- Definition (en francais) -- >
<!ELEMENT DEFENG       - - (%un.txt;)  -- Definition (en anglais) -- >
<!ELEMENT INDEXFRA     - - (%un.txt;)  -- Indexation (en francais) -- >
<!ELEMENT INDEXENG     - - (%un.txt;)  -- Indexation (en anglais) -- >

<!ELEMENT AUTEXP       - - (AUTEXPFR,AUTEXPEN?)  -- Autre(s) expertise(s) -- >
<!ELEMENT AUTEXPFR     - - (%un.txt;)    -- Autre(s) expertise(s) (en francais) -- >
<!ELEMENT AUTEXPEN     - - (%un.txt;)  -- Autre(s) expertise(s) (en anglais) -- >

<!ELEMENT SECTEUR      - - (DOMFRA?,DOMENG?,INDEXFRA*,INDEXENG*)    --
Secteur -- >
<!ELEMENT DOMFRA       - - (%un.txt;)  -- Domaine (en francais) -- >
<!ELEMENT DOMENG       - - (%un.txt;)  -- Domaine (en anglais) -- >

<!ELEMENT DIPLOME      - - (INTITULE,LIEU?,DATE?)  -- Diplome(s) -- >
<!ELEMENT INTITULE     - - (%un.txt;)    -- Intitule du diplome -- >
<!ELEMENT LIEU         - - (%un.txt;)  -- Lieu d'obtention du diplome -- >
<!ELEMENT DATE         - - (%un.txt;)  -- Annee d'obtention du diplome -- >

<!ELEMENT EMPLOI       - -
(PERIODE,NOMORGAN,LIEU?,PAYS,QUALITE?,DESFRA?,DESENG?)  -- Activites
anterieures -- >
<!ELEMENT PERIODE      - - (%un.txt;)    -- Periode -- >
<!ELEMENT QUALITE      - - (%un.txt;)  -- Qualite -- >
<!ELEMENT DESFRA       - - (%un.txt;)  -- Description de l'activite (en francais) -- >
<!ELEMENT DESENG       - - (%un.txt;)  -- Description de l'activite (en anglais) -- >

<!ELEMENT PUBLI        - -
(TYPE,TITRE,LTITRE?,ALTITRE?,LALTITRE?,COAUT?,DATE,INFPUB?)  -- Publications --
>
<!ELEMENT TYPE         - - (%un.txt;)    -- Type de publications -- >
<!ELEMENT TITRE        - - (%un.txt;)    -- Titre original de l'objet de publication -- >
<!ELEMENT LTITRE       - - (%un.txt;)    -- Langue du Titre original de l'objet de publication
-- >
<!ELEMENT ALTITRE      - - (%un.txt;) -- Titre traduit -- >
<!ELEMENT LALTITRE     - - (%un.txt;) -- Langue du Titre traduit -- >
<!ELEMENT COAUT        - - (%un.txt;)  -- Auteurs -- >
<!ELEMENT INFPUB       - - (%un.txt;)  -- Autre(s) information(s) -- >

<!ELEMENT BREVET       - -
(TITRE,INVENT*,NOBREV,DATE,PAYSCOUV?,DEPOSANT?)  -- Brevet -- >
<!ELEMENT INVENT       - - (%un.txt;)  -- Co inventeur(s) -- >
<!ELEMENT NOBREV       - - (%un.txt;)    -- Nø de brevet/Nø de demande -- >
<!ELEMENT PAYSCOUV     - - (%un.txt;)  -- Pays couvert(s) -- >
<!ELEMENT DEPOSANT     - - (%un.txt;)  -- Déposant(s) du brevet -- >

<!ELEMENT DISTSCI      - - (%un.txt;)  -- Distinctions scientifiques,Activites associatives
ou officielles -- >

<!ELEMENT COOPE        - - (INTITULE,NOMORGAN,LIEU,PAYS,DUREE?,DATE?)  --
Cooperations, partenariats ou missions -- >
<!ELEMENT DUREE        - - (%un.txt;)    -- Duree -- >


<!ELEMENT ENSEIGN      - - (COURS*,ENCADR*)  -- Enseignement -- >
```

```
<!ELEMENT COURS        - - (INTITULE,LIEU,DUREE?,DATE?)  -- Cours dispense -- >

<!ELEMENT ENCADR       - - (TITTHES,LIEU?,DATESOUT?)  -- Encadrement de theses --
>
<!ELEMENT TITTHES      - - (%un.txt;)     -- Titre de la these -- >
<!ELEMENT DATESOUT     - - (%un.txt;)  -- Annee de soutenance -- >

<!ELEMENT IST          - - (%un.txt;)  -- Activites de communication (IST) -- >

<!ELEMENT LANGUES      - - (LANGUE,LU?,ECRIT?,PARLE?)     -- Competences
linguistiques -- >
<!-- **1 lu ecrit parle sont facultatifs -->
<!ELEMENT LANGUE       - - (%un.txt;)     -- Langue -- >
<!ELEMENT LU           - - (%m.niv;)      -- Lu -- >
<!ELEMENT ECRIT        - - (%m.niv;)   -- Ecrit -- >
<!ELEMENT PARLE        - - (%m.niv;)   -- Parle -- >

<!ELEMENT CONTACT      - -
(CIVIL,NOM,PRENOM,TITREID?,FONCTION,NOMETAB,RUE,CODPOST,VILLE,CEDEX?,
                  PAYS,TEL?,FAX?,EMAIL?)     -- Personne a contacter -- >

<!ELEMENT GESTION      - - (DATEENTR,DATEMAJ,NOREF)     -- Gestion de la notice --
>
<!ELEMENT DATEENTR     - - (%un.txt;)     -- Date de la premiere entree FINES -- >
<!ELEMENT DATEMAJ      - - (%un.txt;)     -- Date de la derniere mise a jour FINES -- >
<!ELEMENT NOREF        - - (%un.txt;)     -- Nø de reference de la notice FINES -- >

]>
```

```
<FINES>
    <RECORD>
        <IDENTITE>
            <CIVIL>
            </CIVIL>
            <NOM>
            </NOM>
            <PRENOM>
            </PRENOM>
            <TITREID>
            </TITREID>
            <FONCTION>
            </FONCTION>
        </IDENTITE>
        <ADRESSE>
            <NOMETAB>
            </NOMETAB>
            <RUE>
            </RUE>
            <CODPOST>
            </CODPOST>
            <VILLE>
            </VILLE>
            <CEDEX>
            </CEDEX>
            <PAYS>
            </PAYS>
            <TEL>
            </TEL>
            <FAX>
            </FAX>
            <EMAIL>
            </EMAIL>
        </ADRESSE>
        <ORGAN>
            <NOMORGAN>
            </NOMORGAN>
            <SIGLE>
            </SIGLE>
            <CODE>
            </CODE>
            <UNIRECH>
            </UNIRECH>
        </ORGAN>
        <ACTIVITE>
            <DEFFRA>
            </DEFFRA>
            <DEFENG>
            </DEFENG>
            <INDEXFRA>
            </INDEXFRA>
            <INDEXENG>
            </INDEXENG>
        </ACTIVITE>
        <AUTEXP>
            <AUTEXPFR>
            </AUTEXPFR>
            <AUTEXPEN>
            </AUTEXPEN>
        </AUTEXP>
        <SECTEUR>
```

```
        <DOMFRA>
        </DOMFRA>
        <DOMENG>
        </DOMENG>
        <INDEXFRA>
        </INDEXFRA>
        <INDEXENG>
        </INDEXENG>
</SECTEUR>
<DIPLOME>
        <INTITULE>
        </INTITULE>
        <LIEU>
        </LIEU>
        <DATE>
        </DATE>
</DIPLOME>
<EMPLOI>
        <PERIODE>
        </PERIODE>
        <NOMORGAN>
        </NOMORGAN>
        <LIEU>
        </LIEU>
        <PAYS>
        </PAYS>
        <QUALITE>
        </QUALITE>
        <DESFRA>
        </DESFRA>
        <DESENG>
        </DESENG>
</EMPLOI>
<PUBLI>
        <TYPE>
        </TYPE>
        <TITRE>
        </TITRE>
        <LTITRE>
        </LTITRE>
        <ALTITRE>
        </ALTITRE>
        <LALTITRE>
        </LALTITRE>
        <COAUT>
        </COAUT>
        <DATE>
        </DATE>
        <INFPUB>
        </INFPUB>
</PUBLI>
<BREVET>
        <TITRE>
        </TITRE>
        <INVENT>
        </INVENT>
        <NOBREV>
        </NOBREV>
        <DATE>
        </DATE>
        <PAYSCOUV>
```

```
        </PAYSCOUV>
        <DEPOSANT>
        </DEPOSANT>
</BREVET>
<DISTSCI>
</DISTSCI>
<COOPE>
        <INTITULE>
        </INTITULE>
        <NOMORGAN>
        </NOMORGAN>
        <LIEU>
        </LIEU>
        <PAYS>
        </PAYS>
        <DUREE>
        </DUREE>
        <DATE>
        </DATE>
</COOPE>
<ENSEIGN>
        <COURS>
                <INTITULE>
                </INTITULE>
                <LIEU>
                </LIEU>
                <DUREE>
                </DUREE>
                <DATE>
                </DATE>
        </COURS>
        <ENCADR>
                <TITTHES>
                </TITTHES>
                <LIEU>
                </LIEU>
                <DATESOUT>
                </DATESOUT>
        </ENCADR>
</ENSEIGN>
<IST>
</IST>
<LANGUES>
        <LANGUE>
        </LANGUE>
        <LU>
        </LU>
        <ECRIT>
        </ECRIT>
        <PARLE>
        </PARLE>
</LANGUES>
<CONTACT>
        <CIVIL>
        </CIVIL>
        <NOM>
        </NOM>
        <PRENOM>
        </PRENOM>
        <TITREID>
        </TITREID>
```

```
                <FONCTION>
                </FONCTION>
                <NOMETAB>
                </NOMETAB>
                <RUE>
                </RUE>
                <CODPOST>
                </CODPOST>
                <VILLE>
                </VILLE>
                <CEDEX>
                </CEDEX>
                <PAYS>
                </PAYS>
                <TEL>
                </TEL>
                <FAX>
                </FAX>
                <EMAIL>
                </EMAIL>
        </CONTACT>
        <GESTION>
                <DATEENTR>
                </DATEENTR>
                <DATEMAJ>
                </DATEMAJ>
                <NOREF>
                </NOREF>
        </GESTION>
    </RECORD>
</FINES>
```

DBD FINES

1. Description des champs

Labels et indicatifs portent le même nom.

NOM	CONTENU	NATURE	LEXIQUE
NOREF	N° unique de CV attribué par chaque organisme.	Référence	
ACT_DEF	Champs actuels d'expertise ou de compétençe en français et en anglais.	Texte intégral	Oui
ACT_IND	Mots-clés décrivant l'activité (français et anglais).	Liste alphanumérique (décomposition en mots simples)	Oui
ADRESSE	Adresse professionnelle : Nom établissement, N° rue, boite postale, code postal, ville, cedex ou secteur postal, pays, tél, fax, email.	Texte simple	Oui
AUTEXP	Expertises autres que l'actuelle, en frnçais et en anglais.	Texte intégral	Oui
BREVET	Les 3 brevets les plus caractéristiques de l'expertise : Titre, co-inventeurs,N° de brevet, année de dépôt, pays couvert, déposant.	Texte simple	Oui
COD_LAB	Code du laboratoire ou de l'unité de recherche dans l'organisme.	Liste alphanumérique	Oui
CONTACT	Personne à contacter : Titre civil, nom, prénom, tire professionnel, fonction,nom de l'établissement, rue, code postal, ville, cedex ou SP, pays, tél, fax, email.	Texte simple	Oui
COOPE	Coopérations, partenariats : Intiltulé, nom organisme, lieu, pays, durée, date.	Texte simple	Oui
COURS	Cours dispensés : Intitulé, lieu, durée, date.	Texte intégral	Oui
DAT_DIPL	Année d'obtention du dipôme. (Zone inutilisée dans un premier temps)	Liste date	Oui
DATEENTR	Date de la première entrée FINES.	Liste date.	Oui
DATEMAJ	Date de la dernière mise à jour.	Liste date.	Oui
DIPLOME	Intitulé, lieu et date d'obtention.	Texte simple	Oui
DISTSCI	Distinctions scientifiques, activités associatives ou officielles.	Texte simple	Non
EMPLOI	Activité antérieures : Période, nom organisme, lieu, pays, qualité, description de l'activité en français et en anglais.	Texte intégral	Oui
ENCADR	Encadrement de thèses : Titre thèse, lieu, date de soutenance.	Texte intégral.	Oui
IDENTITE	Identité de l'expert : Titre civil, nom, prénom, titre professionnel, fonction.	Texte intégral.	Oui
IST	Activité de communication.	Texte simple.	Oui
LANGUES	Compétences linguistiques avec, pour chaque langue, le niveau de compétence (lu, parlé, écrit : 0 à 3).	Texte intégral.	Oui
NOMORGAN	Nom de l'organisme duquel dépend l'expert.	Liste alphanumérique. (décomposition en mots simples).	Oui

PUBLI	Publications : Type de publication, titre, langue du titre, titre traduit, auteur et co-auteurs, date parution, autres infos.	Texte intégral.	Oui
SECTEUR	Secteur d'application des compétences : Description en français et en anglais Mots-clés en français et en anglais.	Texte intégral.	Oui
SIG_ORG	Sigle de l'organisme duquel dépend l'expert	Liste alphanumérique. (décomposition en mots simples).	Oui
UNIRECH	Intitulé du laboratoir ou unité de recherche.	Texte intégral.	Oui

Annexe 2

Normalisation des fichiers

BUREAU VAN DIJK

INGENIEURS CONSEILS EN GESTION

NORMALISATION FICHIERS FINES

- Dates d'entrée et de mise à jour attendues sous la forme SSAAMMJJ.

  OBC, 8 caractères numériques,

ou	JJMMSSAA	CIRAD
	SSAA/MM/JJ	INSERM
	SSAA-M-JJ	CNRS
	SSAA-JJ-MM	CNRS
	JJMMAA	INRIA
	JJ/MM/AA	INRA
	Absence	ORSTOM, CNRS

- Présence de ß et ^p sur tags SGML — INSERM
  <^p — CIRAD
- N° de référence identique sur tous les CV — INRIA
- Présence de ß devant tags SGML — INRIA, INRA
- Présence d'un caractère FIN SECTION (ß) — ORSTOM
- Doubles tags <identité> et pas de tags <NOM> — INSERM
- Présence de données non conformes en début de fichier — CNRS
- Présence de ^t inutiles
- Balises <FINES> </FINES< encadrant chaque CV — ORSTOM
- Présence de tags sans donnée — ORSTOM
- Absence de la balise </PUBLI> à la fin d'une occurence <PUBLI> (présence seulement en fin de liste) — INSERM
- Sur certaines descriptions, présence de ß au lieu de <P> </P>

Conversion de fichiers SGML pour le Bureau Van Dijk

Différences de formats

Problèmes liés aux caractères CR/LF (marques de paragraphes)

Les différents fichiers SGML ont chacun leur organisation de paragraphes.

Par exemple, le fichier Cirad possède un champ par ligne, et de ce fait une chaîne CR/LF par ligne :

```
<IDENTITE>
<CIVIL>M.</CIVIL>
<NOM>GOUNEL</NOM>
<PRENOM>Christian</PRENOM>
<FONCTION>Ingénieur cartographe</FONCTION></IDENTITE>
```

Le fichier Inserm, au contraire, place chaque champ à la queue leuleu.

```
<record><identite><civil>M.</civil><nom>AUNIS</nom><prenom>DOMINIQUE</prenom><foncti
on>DR1</fonction></identite><adresse><nometab>CENTRE DE NEUROCHIMIE DU
CNRS</nometab><rue>5, RUE BLAISE
PASCAL</rue><codpost>67084</codpost><ville>STRASBOURG</ville><cedex>CEDEX</cedex
>
```

La chaîne CR/LF est bizarrement totalement exclue de ce fichier.

Problèmes liés aux tabulations

Le fichier Inria01 possède une organisation plus simple à visualiser, mais il est trop différent des autres formats. En effet, des caractères de tabulation mettent en évidence les parties et les sous-parties, faisant ainsi apparaître une certaine hiérarchie.

```
<record>
    <identite>
        <CIVIL>MME.</CIVIL>
        <NOM>COQUILLART</NOM>
        <PRENOM>Sabine</PRENOM>
        <FONCTION>Chargé de Recherche</FONCTION>
    </identite>
```

Malheureusement, cette organisation n'existe que dans ce fichier et peut créer des incompatibilités avec les autres SGML.

Problèmes liés aux normes des tables de caractères

Le fichier Orstl_34, en plus des mêmes problèmes que le fichier Inserm, a été enregistré dans un standard en voie de disparition appelé OEM, qui correspond au format ASCII du DOS.

```
<FINES><RECORD><IDENTITE><CIVIL>M.</CIVIL><NOM>CHARVIS</NOM><PRENOM>PHI
LIPPE</PRENOM><TITREID>Docteur de sp,cialit, (g,ophysique)</TITREID><FONCTION>Charg,
de recherche de premiŠre classe</FONCTION></IDENTITE><ADRESSE><NOMETAB>Universit,
```

Paris VI, Laboratoire de g,odynamique sous-marine, Observatoire oc,anologique.</NOMETAB><RUE>BP

Tous les caractères non alphabétiques, commes les caractères accentués et les caractères spéciaux, sont remplacés par d'autres caractères qui rendent ce texte impropre à l'interrogation.

Ce problème se résoud facilement par une simple conversion, à condition d'avoir pris connaissance du format. En effet, la détection automatique d'un standard OEM par rapport à l'ANSI est très hasardeuse.

Format idéal

Il n'existe pas de format idéal, pourvu que tout le monde s'entende sur un format unique de fichier SGML. En effet, puisque chaque fichier que nous avons en exemple a engendré des modifications importantes dans l'écriture du programme de conversion, tout autre type de fichier est susceptible de créer des problèmes que nous ne pouvions pas prévoir.

La macro

La macro suivante ouvre tous les fichiers spécifiés dans Listefic.txt et effectue sur chacun d'eux toutes les opérations de Rechercher/Remplacer spécifiées dans Retr3.txt.

La macro a été réalisée en WordBasic version 6 et fonctionne quel que soit le fichier SGML fourni en exemple.

Texte de la macro

```
Sub MAIN
ListeFic$ = "C:\bases\vandijk\travail\LISTEFIC.TXT"
ListeRetR$ = "C:\bases\vandijk\travail\RETR3.TXT"
Dim F$
On Error Goto GestErreur
Open ListeFic$ For Input As #1
While Not Eof(#1)
    Read #1, NomFic$
    FichierOuvrir .Nom = NomFic$
    Open ListeRetR$ For Input As #2
    While Not Eof(#2)
        Read #2, F$, R$
        EditionRemplacer .Rechercher = F$, .Remplacer = R$, \
        .MotEntier = 0, .RemplacerTout, .RespecterLaCasse = 0
    Wend
    Close #2
    REM FichierFermer 1
Wend
Goto Fin
GestErreur:
MsgBox "Erreur d'exécution. Vérifiez les fichiers modifiés !", 16
Fin:
End Sub
```

Listefic.txt

```
c:\bases\vandijk\travail\cirad.sgm
c:\bases\vandijk\travail\inria01.sgm
c:\bases\vandijk\travail\inserm.sgm
c:\bases\vandijk\travail\orst1_34.sgm
```

Retr3.txt

```
@,&
^p^t^t^t^t,^p
^p^t^t^t,^p
^p^t^t,^p
^p^t,^p
</ACTIVITE>,^p
</ADRESSE>,^p
</ALTITRE>,^p
</AUTEXP>,^p
</AUTEXPFR>,^p
</AUTEXPEN>,^p
</BREVET>,^p
</CEDEX>,^p
</CIVIL>,
</COAUT>,^p
</CODE>,^p
</CODPOST>,
</CONTACT>,^p
</COOPE>,^p
</COURS>,^p
</DATE>,^p
</DATEENTR>,^p
</DATEMAJ>,^p
</DATESOUT>,^p
</DEFFRA>,^p
</DEFENG>,^p
</DEPOSANT>,^p
</DESFRA>,^p
</DESENG>,^p
</DIPLOME>,^p
</DISTSCI>,^p
</DOMFRA>,^p
</DOMENG>,^p
</ECRIT>,^p
</EMAIL>,^p
</EMPLOI>,^p
</ENCADR>,^p
</ENSEIGN>,^p
</FAX>,^p
</FONCTION>,^p
</GESTION>,^p
</IDENTITE>,^p
</INDEXFRA >,
</INDEXFRA>,^p
</INDEXENG >,
</INDEXENG>,^p
</INFPUB>,^p
</INTITULE>,^p
</INVENT>,^p
</IST>,^p
</LALTITRE>,^p
</LANGUE>,^p
```

</LANGUES>,^p
</LIEU>,^p
</LTITRE>,^p
</LU>,^p
</NOBREV>,^p
</NOM>,
</NOMETAB>,^p
</NOMORGAN>,^p
</NOREF>,^p
</ORGAN>,^p
</P>,^p
</PARLE>,^p
</PAYS>,^p
</PAYSCOUV>,^p
</PERIODE>,^p
</PRENOM>, ;
</PUBLI>,^p
</QUALITE>,^p
</RECORD>,^p@d^p
</RUE>,^p
</SECTEUR>,^p
</SIGLE>,^p
</TEL>,^p
</TITRE>,^p
</TITREID>, ;
</TYPE>,^p
</TITTHES>,^p
</UNIRECH>,^p
</VILLE>,^p
</DUREE>,^p
</FINES>,^p
<ACTIVITE>,_____^pActivité : ^p
<ADRESSE>,_____^pAdresse : ^p
<ALTITRE>,Titre traduit :
<AUTEXP>,_____^pAutres expertises : ^p
<AUTEXPFR>,Autre expertise :
<AUTEXPEN>,Autre expertise :
<BREVET>,_____^pBrevet : ^p
<CEDEX>,
<CIVIL>,
<COAUT>,Auteur(s) :
<CODE>,Code :
<CODPOST>,
<CONTACT>,_____^pContact : ^p
<COOPE>,_____^pCoopérations, partenariats ou missions : ^p
<COURS>,Cours dispensé :
<DATE>,Date :
<DATEENTR>,Date de la première entrée Fines :
<DATEMAJ>,Date de la dernière mise à jour Fines :
<DATESOUT>,Soutenue en
<DEFFRA>,Définition :
<DEFENG>,Définition :
<DEPOSANT>,Déposants :
<DESFRA>,Description de l'activité :
<DESENG>,Description de l'activité :
<DIPLOME>,_____^pDiplôme : ^p
<DISTSCI>,Distinctions scientifiques : ^p
<DOMFRA>,Domaine :
<DOMENG>,Domaine :
<DUREE>,Durée :

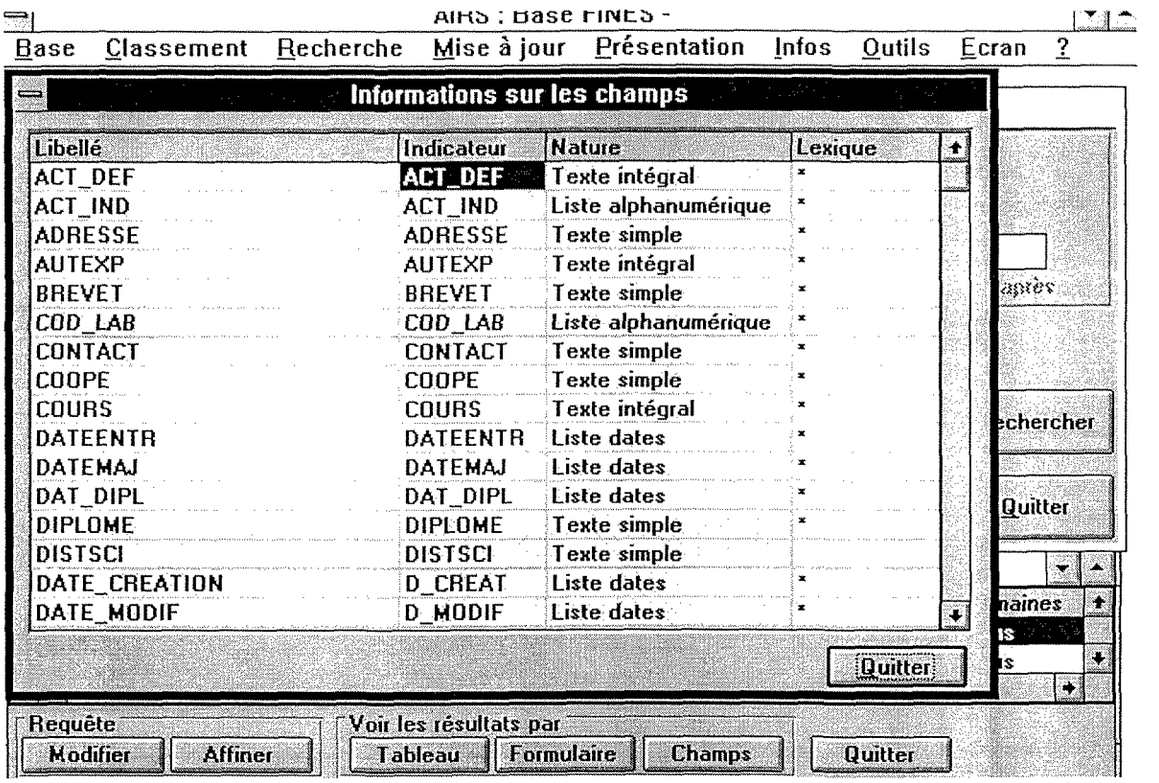
<ECRIT>,Ecrit :
<EMAIL>,Adresse Email :
<EMPLOI>,_____^pEmploi : ^p
<ENCADR>,_____^pEncadrement de thèse : ^p
<ENSEIGN>,_____^pEnseignement : ^p
<FAX>,Fax :
<FINES>,@t Fines^p
<FONCTION>,(FonctionAF)
<GESTION>,_____^pGestion de la notice : ^p
<IDENTITE>,
<INDEXFRA >, ;
<INDEXFRA>,Index :
<INDEXENG >, ;
<INDEXENG>,Index :
<INFPUB>,
<INTITULE>,Intitulé :
<INVENT>,Inventeur(s) :
<IST>,Activités de communication :
<LALTITRE>,Langue du titre traduit :
<LANGUE>,Langue :
<LANGUES>,_____^pLangues : ^p
<LIEU>,Lieu :
<LTITRE>,Langue :
<LU>,Lu :
<NOBREV>,Numéro de brevet ou de demande :
<NOM>,(NomAF)
<NOMETAB>,
<NOMORGAN>,Nom de l'organisme :
<NOREF>,Numéro de référence de la notice Fines :
<ORGAN>,_____^pOrganisme : ^p
<P>,
<PARLE>,Parlé :
<PAYS>,
<PAYSCOUV>,Pays couverts :
<PERIODE>,Période :
<PRENOM>,(PrénomAF)
<PUBLI>,_____^pPublication : ^p
<QUALITE>,Qualité :
<RECORD>,^p@d Curriculum Vitae^p
<RUE>,
<SECTEUR>,_____^pSecteur : ^p
<SIGLE>,Sigle :
<TEL>,Téléphone :
<TITRE>,Titre :
<TITREID>
<TITTHES>,Titre de la thèse :
<TYPE>,Type
<UNIRECH>,
<VILLE>,
^t^t^p,^p
^t^p,^p
^p,^p
^p,^p
^t^p,^p
^p^p^pIndex : ,^pIndex :
^p ^pIndex : ,^pIndex :
^p^pIndex : ,^pIndex :
^p^p^p_____,^p_____
^p^p_____,^p_____
^p^p^pNom : ,^pNom :

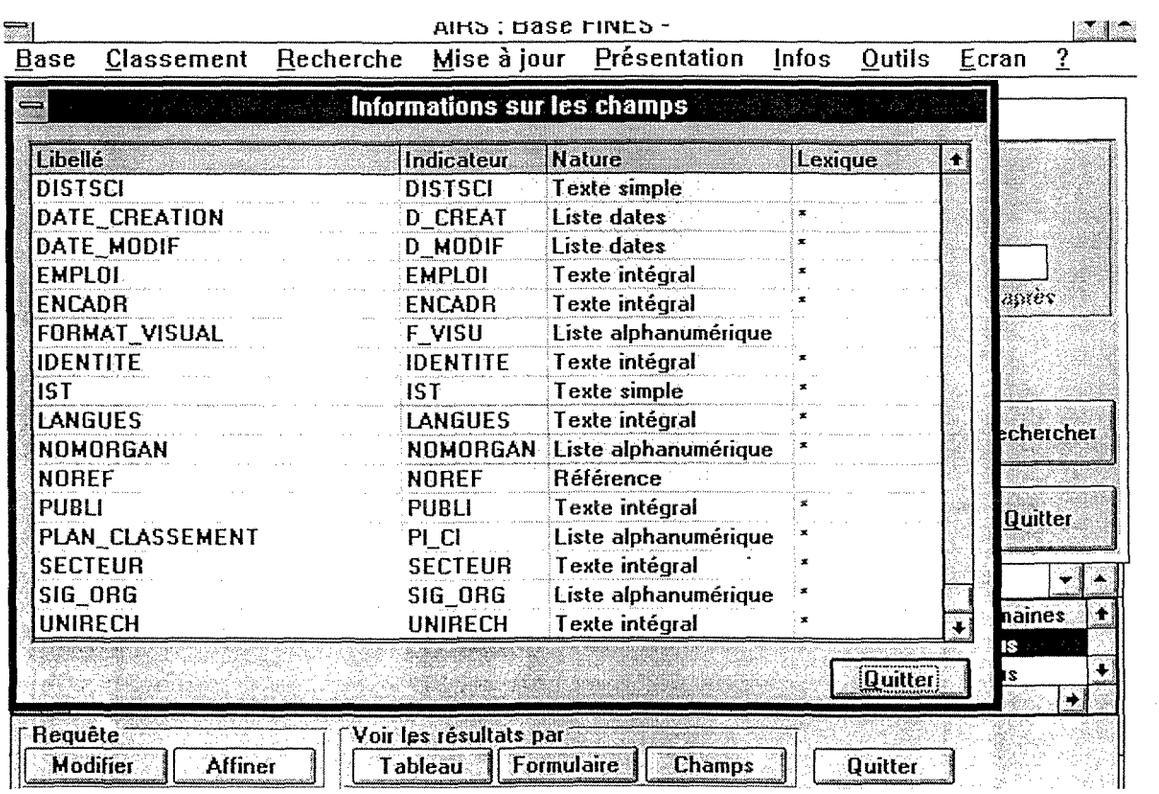
^p^pNom : ,^pNom :
^p(NomAF),
(NomAF),
^p(PrénomAF),
(PrénomAF),
^p(FonctionAF),
(FonctionAF),
FRA^p,France^p
^p^p^p^p,^p
^p^p^p,^p
^p^p,^p
^p^p,^p
^p @d,^p@d
^p@d Curriculum Vitae,^p^p@d Curriculum Vitae

Annexe 3

Copies d'écran

Liste des champs





Formulaire de recherche

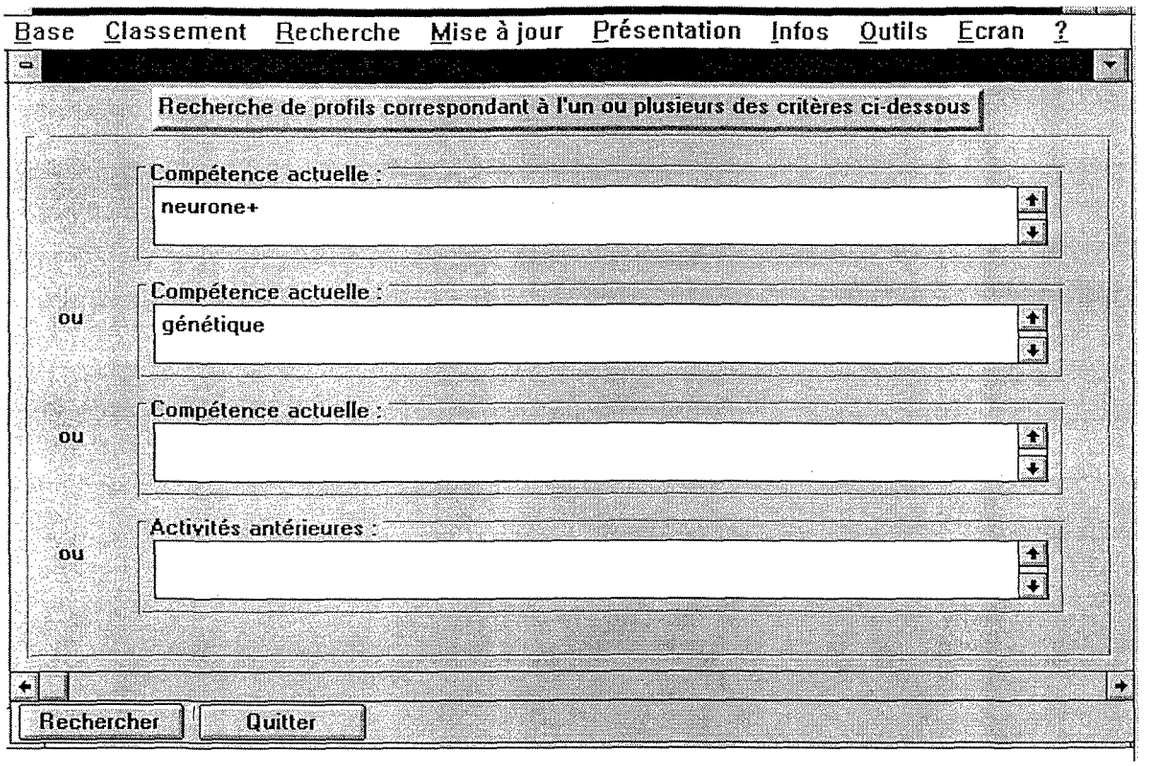
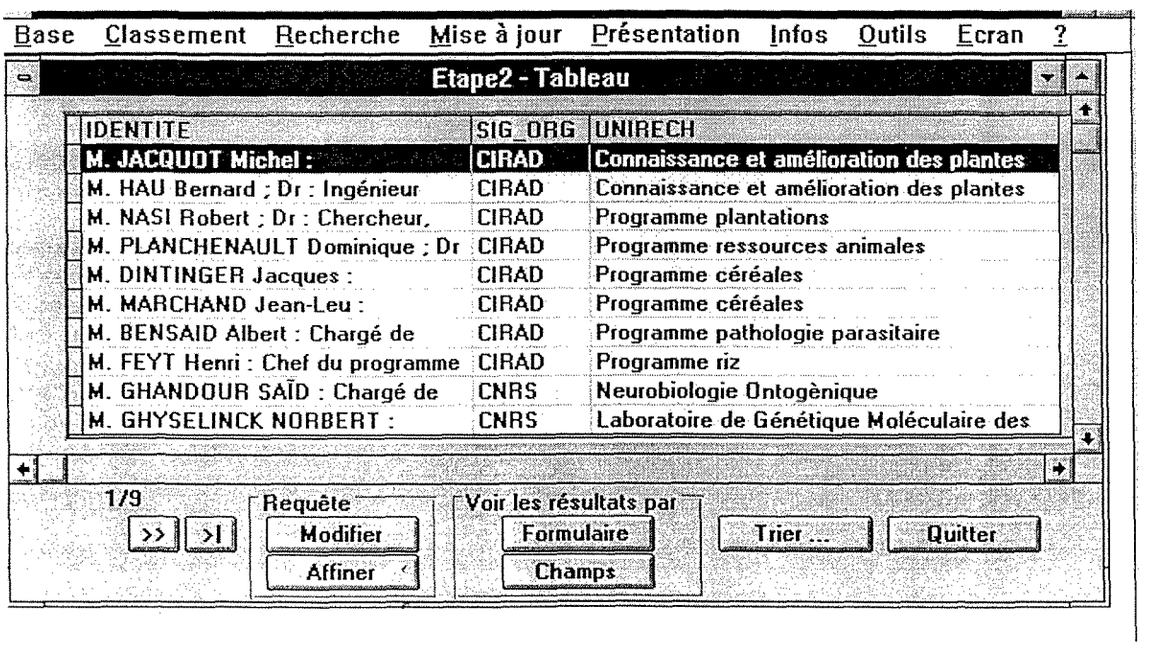
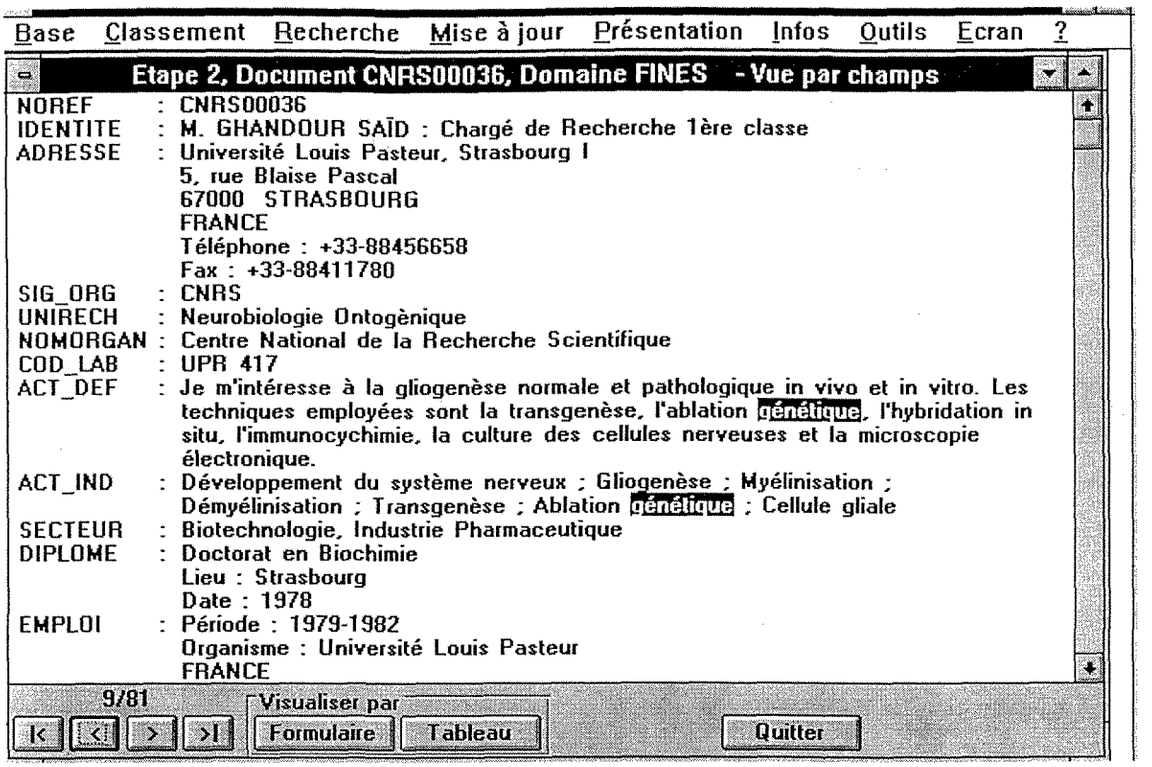


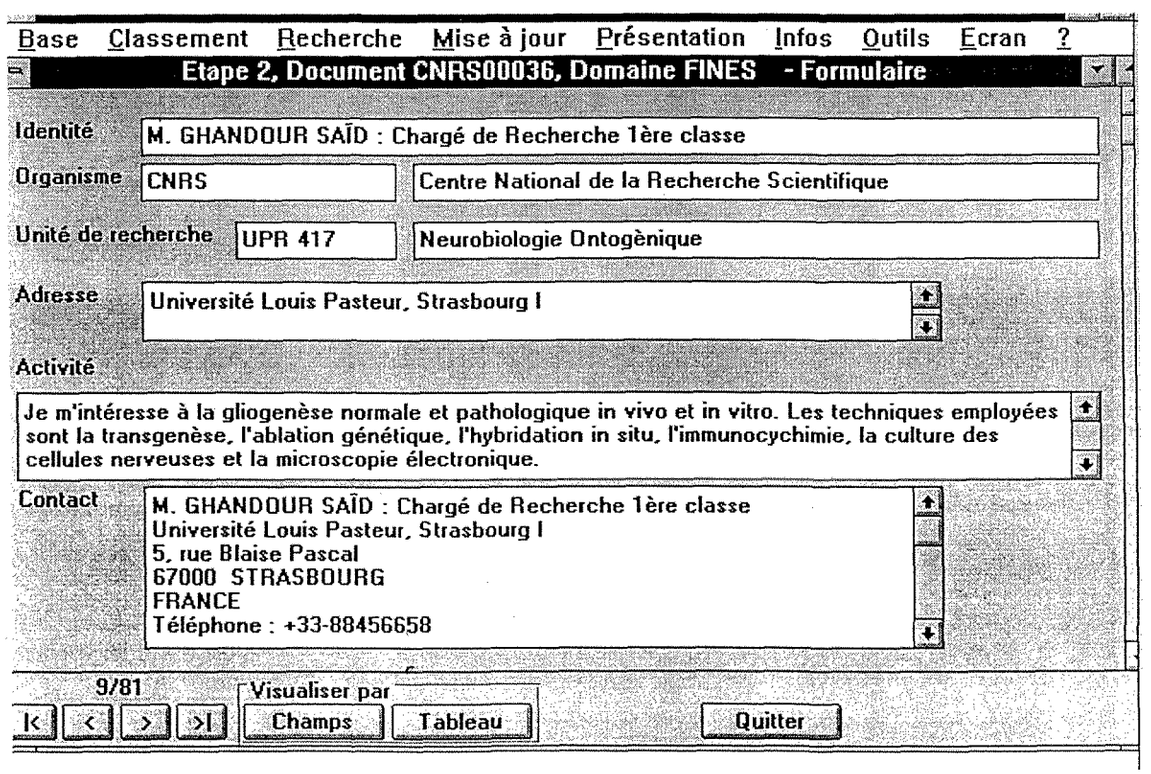
Tableau de résultat



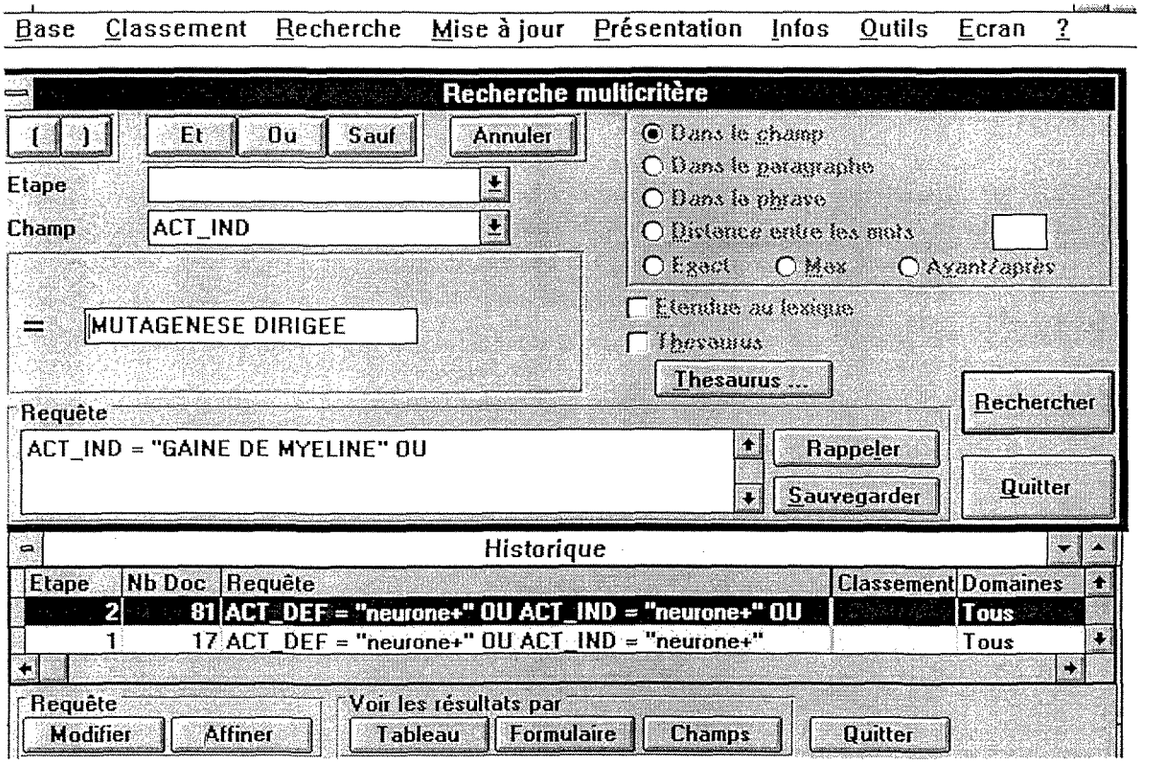
Vue par champs



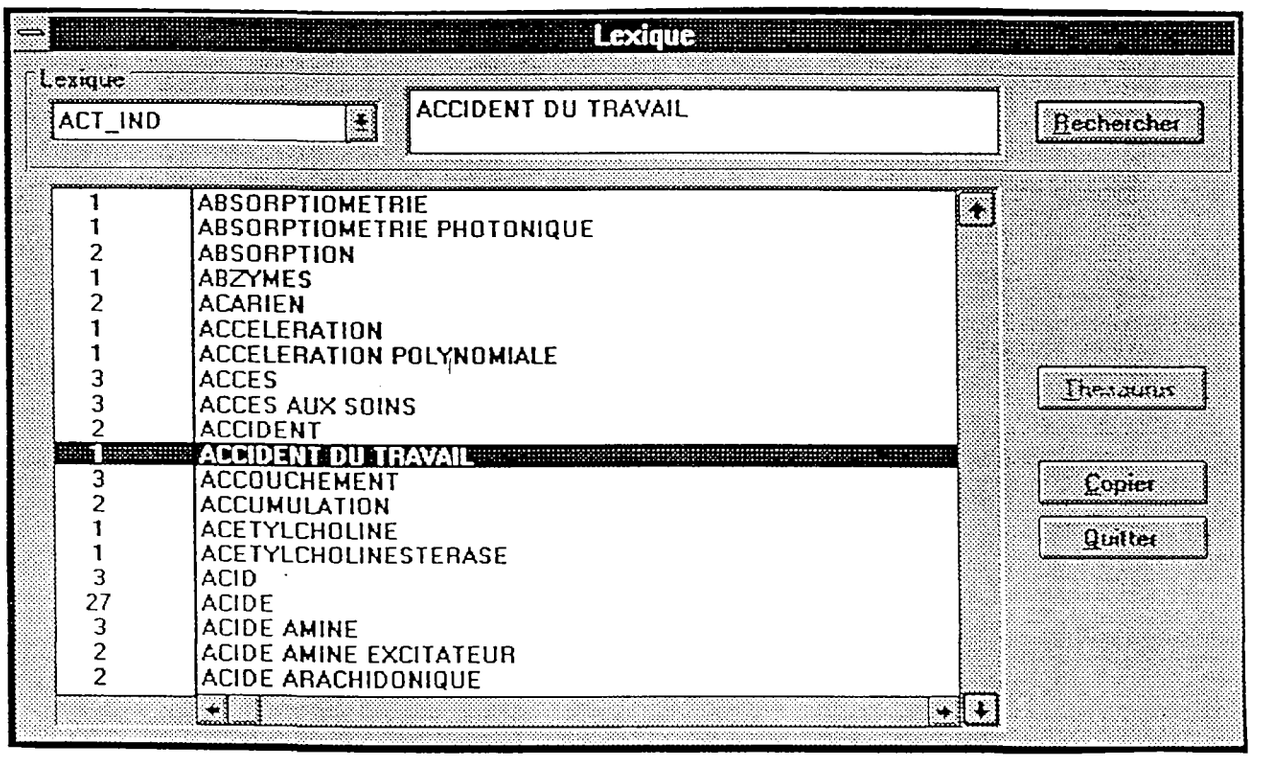
Vue par formulaire



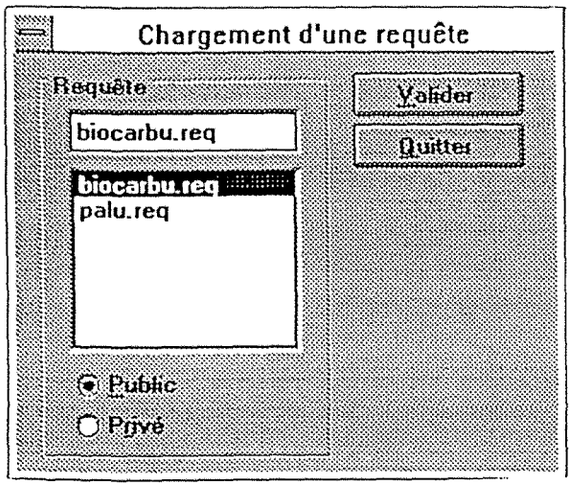
Recherche multicritère



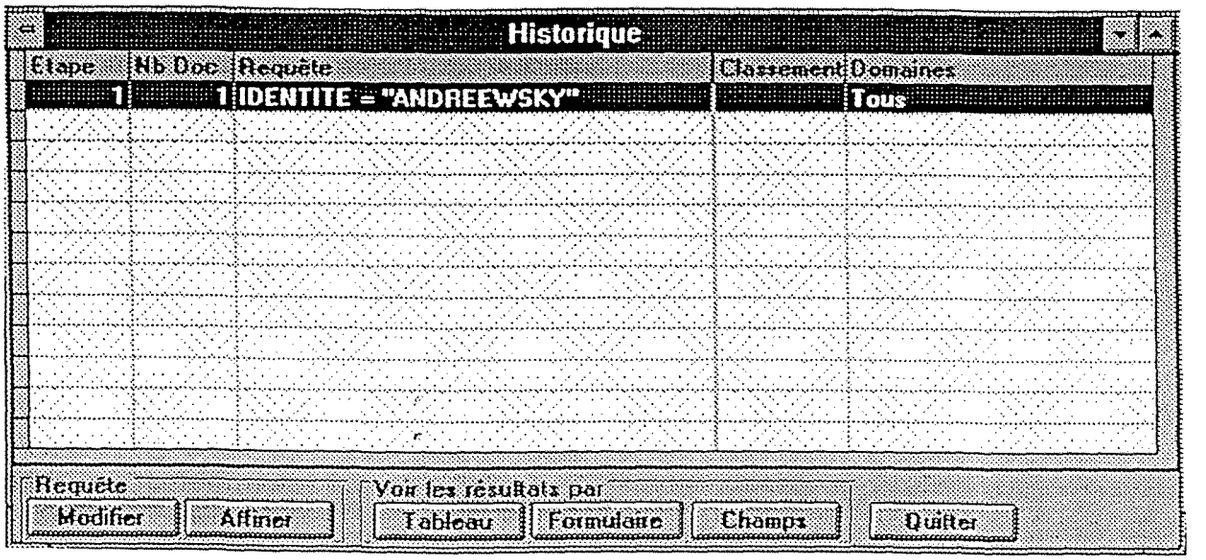
LEXIQUE


Lexique
Lexique
ACT_IND
ACCIDENT DU TRAVAIL
Rechercher
1 ABSORPTIOMETRIE
1 ABSORPTIOMETRIE PHOTONIQUE
2 ABSORPTION
1 ABZYMES
2 ACARIEN
1 ACCELERATION
1 ACCELERATION POLYNOMIALE
3 ACCES
3 ACCES AUX SOINS
2 ACCIDENT
1 ACCIDENT DU TRAVAIL
3 ACCOUCHEMENT
2 ACCUMULATION
1 ACETYLCHOLINE
1 ACETYLCHOLINESTERASE
3 ACID
27 ACIDE
3 ACIDE AMINE
2 ACIDE AMINE EXCITATEUR
2 ACIDE ARACHIDONIQUE
Thesaurus
Copier
Quitter

SAUVEGARDE/RAPPEL DES REQUETES



HISTORIQUE DES RECHERCHES



Annexe 4

Grilles d'évaluation

GRILLE D'EVALUATION FINES

Nombre de consultation faites : .....

Nombre de consultations :

- à l'aide du formulaire : .....
- en recherche multicritère : .....

Champs les plus utilisés dans les requêtes : .....

Nombre moyen de termes utilisés dans une requête : .....

Nombre moyen de CV en réponse à une enquête : .....
(autre qu'une recherche sur un nom ou un organisme)

Évaluation du nombre de réponses non pertinentes en réponse à une requête : .....

.....

Raisons du bruit observé :

- imprécision du vocabulaire de description des activités : .....
- indexation trop complexe : .....
- absence de hiérarchisation : .....
- autres : .....

.....

Avez-vous pu mettre en évidence du silence dans les réponses ?
Si oui, taux de silence : .....

.....

Raison constatées du bruit :

- indexation absente ou trop large : .....
- problèmes de synonymie : .....
- absence de hiérarchisation : .....
- autres : .....

.....

La qualité de la saisie des données a t-elle été pour vous sources d'erreurs ?
Si oui, types d'erreurs constatés : .....

.....

.....

L'indexation vous a t-elle parue :

- trop fine : .....
- trop large : .....
- bonne : .....
- imprécise : .....
- cohérente d'un CV à l'autre : .....
- cohérente d'un organisme à l'autre : .....

D'après vous, faudrait-il donner des règles plus précises dans la rédaction des CV :

- précision du vocabulaire : .....
- niveau de détail dans les champs : .....
- volume des données dans les rubriques : .....
- etc... : .....

.....

Pour vous l'absence d'un vocabulaire contrôlé est-il gênant : .....

.....

Quels sont les types de recherche que vous auriez souhaités faire :

- recherche géographique : .....
- autre : .....

.....

.....

A l'usage, quelles sont les remarques que vous avez à formuler sur la structure de la base : .....

.....

.....

- Le niveau de structuration des champs est-il suffisant ou excessif ?

  .....

  .....

- Y a t-il des rubriques qui vous paraissent manquer. Si oui, lesquelles ?

  .....

  .....

- Certaines rubriques vous semblent-elles inadéquates. Si Oui, lesquelles ?

  .....

  .....

- Le contenu des rubriques entre les notices de diverses origines (volume, présentation, style rédactionnel, contenu) vous apparaît-il comme cohérent ?

  .....

  .....

Pouvez-vous donner en annexe des exemples typiques des requêtes faites sur la base ?

- requêtes n'ayant pas abouti : .....

  .....

  .....

- requêtes ayant abouti :

  - avec des résultats pertinents :

    .....

    .....

  - avec des résultats non pertinents :

    .....

    .....

SUPPORT TECHNIQUE POUR L'EVALUATION

Evaluation du logiciel

1. Aspects généraux.

– Qualité du menu d'accueil.

– Existence de différents niveaux de dialogue en fonction des catégories d'utilisateurs. Aide en ligne.

– Niveau général d'ergonomie du logiciel : adaptation du logiciel aux objectifs du projet. Présente-t-il trop ou pas assez de fonctionnalités ?

– Ergonomie de l'interface utilisateur pour l'accès aux fonctions du logiciel.

– Clarté de la présentation des écrans.

– Evaluation du choix des écrans de réponse.

– Facilité d'interrogation : l'accès aux données est-il plus aisé en interrogeant globalement le CV ou champ par champ ? Quelles sont les modalités d'interrogation si on choisit de ne pas interroger par langage libre – ce qui va à l'encontre de la philosophie de la base telle que définie initialement et ne place pas le logiciel dans ses conditions optimales d'utilisation – ? Quels sont les opérateurs booléens admis ? Quels sont les opérateurs de proximité, les opérateurs de troncature et les opérateurs de comparaison admis ?

– Le taux de récupération devrait être déterminé par chaque organisme sur son propre corpus de données.

– Syntaxe des questions en langage libre pour l'analyse contextuelle : quelle syntaxe fournira les meilleurs résultats "maladie fruit zone tropicale" ou "les maladies des fruits en zone tropicale" ?

– Interrogation et jeux de caractères : ce genre de logiciel sait désaccentuer mais pas réaccentuer. Obtient-on des réponses équivalentes selon que l'on interroge par "INGENIEUR", "ingenieur" ou "ingénieur" ?

– Evaluation des fonctionnalités particulières du logiciel :
- bases personnalisées (bases enfants) servant à circonscrire les domaines de recherche.
- dictionnaires de jargon : comment aborder les problèmes posés par la polysémie ? Convient-il de faire différents dictionnaires de jargon spécialisés par domaine ou par organisme ? On risque alors de se trouver face à de graves ambiguïtés, ainsi :
  GED = Gestion Electronique de Documents
  GED = Groupement Européen d'Edition
- historiques (ex : catalogage des questions posées).

– Evaluation des conditions d'accès (TRANSPAC).

2. Exploitation des résultats de la recherche.

– Quelles sont les possibilités de tri, de classement des réponses, peut-on cataloguer la séquence des questions et des réponses ?

– Les différentes modalités choisies pour la présentation des réponses sont-elles satisfaisantes ?

– L'absence de visualisation (par effet vidéo) des termes de la question dans les notices en réponse, du fait du traitement sémantique de la question en langage naturel, est-elle gênante ?

– La recopie d'écran est-elle facile à mettre en œuvre ?

– Est-il intéressant de créer des bases enfants virtuelles à partir des réponses (bases thématiques virtuelles) ? Peut-on ajouter des commentaires à un document, peut-on pondérer les CV ?

3. Sécurité et confidentialité.

– Quelles sont les possibilités de suppression logique d'un document ?
– Dispose-t-on d'une historique des suppressions ?
– Certaines procédures sont-elles protégées ?
– Possibilités d'interdiction de communication de documents ou de parties de documents ?
– Cryptage de données ?
– Existe-t-il des dispositifs de protection et de contrôle d'accès au niveau :
- utilisateurs ?
- documents ou parties de documents ?
- fonctions ?

– Existe-t-il un dispositif d'alerte en cas d'utilisation "anormale" de la base, un dispositif de détection d'accès illicites ?
– Un document en cours de consultation peut-il être modifié ?
– Existe-t-il un dispositif miroir pour la sécurité d'exploitation ?
– En cas d'incident système, quelles sont les sécurités ?

B. Evaluation de la base de données.

– Le niveau actuel de structuration des champs est-il suffisant ? excessif ?
– Le contenu des champs est-il cohérent, pour l'ensemble de la base, dans sa présentation, son style sémantique, son contenu intellectuel ?
– Le degré de remplissage des champs est-il cohérent dans l'ensemble de la base ?

C. Administration de la base.

1. Gestionnaire.

– Est-il possible de tenir un journal de bord d'exploitation (nombre de documents, espace mémoire libre, etc) ?

– A-t-on des possibilités de mesure de l'activité de la base (nombre de documents nouveaux, nombre d'utilisateurs, nombre d'interrogations) ?
– Avis sur les procédures éventuelles de mise à jour notice par notice ou d'écrasement par un nouveau fichier en batch.
– Avis sur les procédures de gestion des rejets, de gestion des doublons.
– Nombre d'utilisateurs possibles en simultané.

2. Utilisateurs.

– Avis sur la procédure de chargement des mises à jour : instaure-t-on une possibilité d'action en ligne par l'utilisateur, notice par notice ?
Si c'est le cas, quel est le degré d'ergonomie de saisie : contrôle des zones obligatoires, duplication, couper / coller ? Quelles sont les possibilités de sauvegarde en cas d'accident ? Existe-t-il un journal des modifications ?
– Facilités de contrôle des notices par les utilisateurs avant envoi (parseur pour alimentation de la base).

Dépannage.

Qui contacter en cas de problèmes :
- d'accès ?
- d'utilisation du logiciel ?
- d'administration de la base ?

Annexe 5

Contenu de la maquette

Contenu de la maquette

Au 21 mai 1996, la maquette contient 689 Curriculum Vitae dont l'origine se décline comme suit :

Organisme	Nombre de CV prévus pour la maquette	Nombre de CV fournis
CEA	200	0
CIRAD	100	101
CNRS	350	189
INRA	150	48
INRIA	50	35
INSERM	200	201
ORSTOM	175	115
Total	1235	689

Remarques :

- Nos interlocuteurs au CEA n'ont, à l'heure actuelle, toujours pas reçu l'autorisation d'alimenter la maquette en données et ce, apparemment, pour des problèmes de confidentialité.

- Il est admis que la base RETIS du CNRS alimentera progressivement FINES.

Annexe 6

Bilan IDT 96

FINES : Bilan d'IDT 96

Voici la synthèse des observations qui ont été faites par les visiteurs sur le stand du MENESR.

- Le projet est jugé intéressant et opportun tant par le privé que le public (Elf, Renault, ADIST, etc.).

- La mise en exploitation devrait démarrer au plus vite.

- Le choix du support de diffusion n'est pas évident : les CD sont chers et difficiles à mettre à jour avec une fréquence suffisante ; Internet intéresse les chercheurs mais beaucoup moins les industriels (pour l'instant). Le Minitel semble plus intéressant à ces derniers mais son avenir est incertain.
  Solution proposée : double diffusion CD et on-line (avec l'accès le plus large possible au départ, pour "lancer" la base : peut-être Minitel avec très peu de données en premier accueil et passage ensuite à Internet si accord organismes concernés et accord CNIL. Il est d'ailleurs possible que, si la base n'est pas disponible avant un an et demi - deux ans, Internet soit alors très répandu, même dans l'Industrie).

- Il serait très intéressant de faire entrer entre les Universités et même le privé dans FINES (opinion émise par un représentant d'Elf-Atochem). Le MENESR n'a pourtant, apparemment, pas eu beaucoup de répondant de la part de ces deux types d'acteurs.

Annexe 7

Comptes rendus de réunions du Comité de Pilotage

MINISTERE DE L'EDUCATION NATIONALE
DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE

DIRECTION DE L'INFORMATION SCIENTIFIQUE,
DES TECHNOLOGIES NOUVELLES
ET DES BIBLIOTHEQUES

Sous-direction de l'Information
Scientifique et Technique

Département de l'Information
Spécialisée

Affaire suivie par : Mme Dominique VALLEE
Téléphone : 46 34 36 31
Télécopie : 46 34 48 48
N/Réf. DISTNB/DV/96/150

Paris, le 28 février 1996

COMPTE RENDU DE LA REUNION DU COMITE DE PILOTAGE FINES DU 20 FEVRIER 1996

Après une introduction de Daniel Confland, Pierre-Gilles Pélisson a fait le point de la situation actuelle du projet et a exposé les buts de la réunion, en l'absence de Dominique Vallée, souffrante.

Etat du projet FINES

Suite au départ de la Société GEDIMAGE du concepteur de LEXIC, il s'est avéré impossible de maintenir le logiciel et la décision a dû être prise de choisir un autre système de gestion et d'interrogation pour la maquette FINES. Le choix s'est opéré avec le Bureau van DIJK et l'accord du MENESR. Entre diverses solutions, le progiciel de gestion documentaire AIRS de la Société EURITIS (European Information Technology for Information Science) a été retenu. Cet outil possède de multiples fonctionnalités : recherche en texte intégral, lexiques, thesaurus, extraction paramétrable, etc. Il permet de stocker et de retrouver des documents en SGML.

AIRS est utilisé pour la gestion documentaire à l'ADEME, à l'UNESCO, à l'INRS, au Ministère de l'Industrie et dans des Centrales nucléaires, entre autres.

Actuellement, la maquette contient 689 notices (CV).

Certains organismes doivent encore nous fournir des notices pour arriver au volume prévu d'un millier de notices.

D'autre part, l'insertion de ces données a fait apparaître certaines erreurs d'écritures au sujet desquelles P.-G. Pélisson contactera prochainement chaque organisme de façon à éviter toute intervention manuelle sur les fichiers ultérieurs pour leur rentrée dans la base. Ceci souligne l'importance de bien traiter les données.

Présentation d'AIRS

La présentation du logiciel et de la maquette a été effectuée par Monsieur Gérard Louis, Directeur Technique et Madame Véronique Roland d'EURITIS.

Les capacités du système ont été démontrées en prenant appui sur des CV fournis par les organismes : restrictions d'accès, recherche par formulaire simple ou recherche multicritères, affichage historique, affichage par tableaux, consultation des données principales du profil par formulaire ou consultation par champs (ensemble du CV), utilisation des lexiques, construction de thesaurus pour gérer les synonymies dans ces lexiques, gestion de la proximité à l'interrogation, possibilités de choix des champs à exporter ou imprimer, etc. D'intéressantes possibilités de développement ont été évoquées pour répondre à diverses questions posées par les membres du Comité de Pilotage.

Prochaines actions et planning.

P.-G. Pélisson a ensuite rappelé les actions à mener dans les plus proches délais :

- PGP prendra contact très rapidement (d'ici la semaine 09) avec les divers organismes afin, d'une part de leur communiquer les précautions à suivre pour éviter certains problèmes d'écriture préjudiciables à l'exploitation des données, et d'autre part de recueillir les caractéristiques des configurations disponibles chez chaque organisme.

- La semaine prochaine (09), un essai de connexion au serveur van DIJK sera fait à partir de l'INIST.

- Courant mars, les connexions devront être opérationnelles chez les différents organismes et les formations à l'utilisation d'AIRS débuteront conjointement.

- Les organismes enverront, d'ici début mars (date butoir le 11 mars), leurs conclusions sur la grille d'évaluation déjà proposée. Ces remarques nous serviront à établir un document de synthèse. L'évaluation portera surtout sur la structure et le contenu de la base, en particulier sur la cohérence des données renseignant des rubriques analogues des CV.

- Il conviendra de fixer rapidement la date de la prochaine réunion du Comité de Pilotage (fin mars) ainsi que la date de fin d'évaluation et de remise du rapport d'évaluation.

MINISTERE DE L'EDUCATION NATIONALE
DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE

DIRECTION DE L'INFORMATION SCIENTIFIQUE,
DES TECHNOLOGIES NOUVELLES
ET DES BIBLIOTHEQUES

Sous-direction de l'Information
Scientifique et Technique

Département de l'Information
Spécialisée

Affaire suivie par : Mme Dominique VALLEE
Téléphone : 46 34 36 35
Télécopie : 46 34 48 58
N/Réf. DISTNB/DV.AP/96/397

INIST - CNRS
ARRIVÉE
26 JUIN 1996
DIRECTION TECHNIQUE

Paris, le 18 jiuin 1996

Compte rendu du Comité de Pilotage FINES du 7 juin 1996

En présence de :

P.G. Pélisson	INIST/CNRS
O. Artur	INIST/CNRS
J. A. Ville	ORSTOM
L. Porges	ORSTOM
C. Bouviala	INSERM
C. Trécourt	INSERM
J. F. Foucher	CIRAD
H. Pampouille	INRA
S. Prévosteau	BRGM
A. Pretet	INRETS
J. Millet	CNRS/SOSI
L. Pierron	INRIA
D. LOHNER	CEA/DIST
M. T. Frèrejacque	CNRS
M. Dejean	Bureau Van Dijk
C. Oto	Bureau Van Dijk

La maquette FINES a été testée par l'ensemble des organismes et démontrée sur le stand du ministère lors du salon IDT. Le tour de table amène les commentaires suivantes :

- l'INRIA a constaté que la base était trop petite pour tester (seulement 35 CV de l'INRIA). Il manque un lexique implicite comportant l'ensemble des termes de recherche. Une certaine hétérogénéité est regrettée dans l'indexation du champ "secteur d'application" ;

- le CEA ajoute aux précédentes remarques la mauvaise adaptation des écrans "formulaires de recherche" et l'absence de verbalisation des intitulés de champ qui gène l'évaluation ;

• le CNRS confirme les points précédents et note qu'il serait opportun d'améliorer la visualisation des résultats ;

• l'INRA constate de plus que le champ "définition activité" est mal rempli et trop hétérogène ;

• le CIRAD a trouvé le logiciel compliqué et peu rapide. Le nombre d'écrans à visualiser avant de poser sa question est trop important. Le CIRAD a néanmoins constaté que la base répond bien aux requêtes et que l'information est intéressante mais trop éclatée et souvent répétée d'un champ à l'autre, d'où la suggestion de simplifier des champs et d'en cumuler certains ;

• l'INSERM confirme les points précédents ; un point positif a été noté qui concerne l'historique et les procédures de modification de la question ;

• l'ORSTOM insiste sur l'hétérogénéité de l'indexation sur certains champs et propose la fusion des champs "domaine d'activité" et "secteur d'application" ;

• l'INIST a fourni un commentaire écrit qui fait la synthèse des réactions entendues à IDT (voir en annexe).

D. Vallée reprend la parole pour indiquer que le ministère est plutôt satisfait de cette première maquette qui montre tout l'intérêt de la future base FINES.

Si l'on résume les commentaires oraux des participants, en attendant leurs commentaires écrits, on note que les améliorations demandées sont :

- la création d'un lexique unique implicite ;
- l'accès par unitermes y compris pour les champs mot clés ;
- la verbalisation des intitulés des champs ;
- l'amélioration des écrans formulaire et visualisation ;
- la création d'un champ numérique sur l'année du diplôme ;
- le cumul éventuel des champs "définition activité" et "secteur application".

Ces points devront être confirmés d'ici fin juin par les partenaires.

D. Vallée indique ensuite que le ministère souhaite que la prochaine réunion soit consacrée à l'examen des conditions de finalisation du projet (aspects techniques, documentaires et de diffusion). Il est important que la base monte maintenant en puissance et que les partenaires continuent à l'alimenter. A ce sujet Mme Frèrejacques pour le CNRS demande qu'un courrier officiel du ministère confirme ce point. En effet, le CNRS n'était engagé que pour la réalisation de la maquette et souhaite être saisi de nouveau pour la suite. L'INSERM approuve cette demande. D. Vallée proposera à M. Dizambourg de signer un courrier en ce sens aux directeurs des organismes de recherche.

Annexe 8

Arrêté au Journal Officiel

Art. 3. – Un délai de trois semaines à compter de la publication du présent arrêté est imparti aux organisations syndicales énumérées à l'article 1er pour désigner leurs représentants.

Art. 4. – La durée du mandat des membres des représentants du personnel prendra fin dans un délai maximum de six mois suivant le renouvellement des commissions administratives paritaires.

Art. 5. – L'arrêté du 18 août 1992 relatif à la liste des organisations syndicales habilitées à désigner des représentants au sein du comité technique paritaire central de l'Institut national de la santé et de la recherche médicale est abrogé.

Art. 6. – Le directeur général de l'Institut national de la santé et de la recherche médicale est chargé de l'exécution du présent arrêté, qui sera publié au *Journal officiel* de la République française.

Fait à Paris, le 25 octobre 1995.

Le ministre de l'éducation nationale, de l'enseignement supérieur, de la recherche et de l'insertion professionnelle,
Pour le ministre et par délégation :
Par empêchement du directeur général de l'administration, des ressources humaines et des affaires financières :
Le chef de service,
J.-F. CERVEL

Le ministre de la santé publique et de l'assurance maladie,
Pour le ministre et par délégation :
Par empêchement du directeur général de la santé :
Le chef de service,
A. LEFEBVRE

Arrêté du 25 octobre 1995 portant création de la maquette d'une application nationale de traitement automatisé d'informations nominatives dénommé Fichier national de l'expertise scientifique réalisée par la direction de l'information scientifique et technique et des bibliothèques

NOR : *MENT9502406A*

Le ministre de l'éducation nationale, de l'enseignement supérieur, de la recherche et de l'insertion professionnelle,

Vu la convention n° 108 du 28 janvier 1981 du Conseil de l'Europe ;

Vu la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, modifiée par la loi n° 88-227 du 11 mars 1988 ;

Vu le décret n° 78-774 du 17 juillet 1978 pris pour l'application des chapitres Ier à IV et VII de la loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, modifié par les décrets n° 78-1223 du 28 décembre 1978, n° 79-421 du 30 mai 1979 et n° 80-1030 du 18 décembre 1980 ;

Vu la lettre de la Commission nationale de l'informatique et des libertés en date du 25 juillet 1995 portant le numéro 39/288,

Arrête :

Art. 1er. – Il est créé au ministère de l'éducation nationale, de l'enseignement supérieur, de la recherche et de l'insertion professionnelle un traitement automatisé d'informations nominatives dénommé Fichier national de l'expertise scientifique (FINES) ; ce traitement a pour objet d'élaborer la maquette d'une banque de données destinée à faire connaître les compétences scientifiques des chercheurs français appartenant aux organismes publics de recherche.

Art. 2. – Les catégories d'informations nominatives sont les suivantes :
- nom, prénoms du chercheur et son sexe ;
- des informations socioprofessionnelles (titre, diplômes, activités de recherche principales et secondaires) ;
- des informations sur les publications et les brevets auxquels le chercheur a participé.

Art. 3. – Les destinataires de ces informations :
- au niveau de l'administration centrale :
  - les agents habilités au sein de la mission scientifique et technique ;
  - les agents habilités à la direction de l'information scientifique et technique et des bibliothèques ;
  - les agents habilités à la direction générale de la recherche et de la technologie ;
- au niveau des organismes de recherche :
  - les agents responsables du projet habilités par le directeur général.

Art. 4. – Le droit d'accès prévu par l'article 34 de la loi du 6 janvier 1978 susvisée s'exerce auprès de l'I.N.I.S.T.-C.N.R.S. (Institut national de l'information scientifique et technique, direction technique des bases de données), 2, avenue du Parc-de-Brabois, 54514 Vandœuvre-lès-Nancy.

Art. 5. – Le directeur général de l'administration, des ressources humaines et des affaires financières et le directeur de l'information scientifique et technique et des bibliothèques sont chargés, chacun en ce qui le concerne, de l'exécution du présent arrêté, qui sera publié au *Journal officiel* de la République française.

Fait à Paris, le 25 octobre 1995.

Pour le ministre et par délégation :
Le directeur du cabinet,
F. DELON

Arrêté du 31 octobre 1995 autorisant au titre de l'année 1996 l'ouverture de concours pour le recrutement d'adjoints administratifs de l'administration centrale (spécialité Administration et dactylographie)

NOR : *MENA9502341A*

Par arrêté du ministre de l'éducation nationale, de l'enseignement supérieur, de la recherche et de l'insertion professionnelle et du ministre de la fonction publique en date du 31 octobre 1995, indépendamment des dispositions législatives et réglementaires relatives aux emplois réservés aux bénéficiaires du code des pensions militaires d'invalidité et des victimes de guerre et aux travailleurs handicapés, est autorisée au titre de l'année 1996 l'ouverture de deux concours pour le recrutement d'adjoints administratifs de l'administration centrale, spécialité Administration et dactylographie.

Le nombre total de places offertes aux concours sera fixé ultérieurement par arrêté.

La date de clôture du registre des inscriptions est fixée au vendredi 1er décembre 1995.

La date des épreuves et la composition du jury feront l'objet d'arrêtés du ministre de l'éducation nationale, de l'enseignement supérieur, de la recherche et de l'insertion professionnelle.

Nota. – Pour tous renseignements, les candidats doivent s'adresser au ministère de l'éducation nationale, de l'enseignement supérieur, de la recherche et de l'insertion professionnelle (direction de l'administration et du personnel, sous-direction des personnels, bureau des concours administratifs), adresse postale : 54, rue de Bellechasse, 75357 Paris 07 SP ; adresse du service : 4, rue Danton, 75006 Paris.