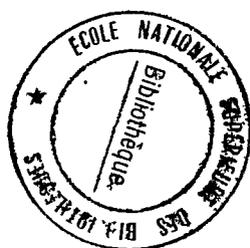


**Ecole Nationale
Supérieure de
Bibliothécaires**

**Université
Claude Bernard
Lyon I**

**DESS Informatique
Documentaire**



Rapport de stage

**CONCEPTION D'UNE BASE DE DONNEES
DOCUMENTAIRE AVEC LE LOGICIEL
MISTRAL**

Hélène BREUL

sous la direction de B. HOURTOULLE
Ingénieur au Commissariat à l'énergie atomique

1990
Stages
7

1990

SOMMAIRE

REMERCIEMENTS	2
INTRODUCTION	3
1ère Partie : PRESENTATION DE L'ENTREPRISE ET DES MOYENS INFORMATIQUES	4
I. LE COMMISSARIAT A L'ENERGIE ATOMIQUE (C.E.A.)	5
II. LES CENTRES	5
III. LA DAM	6
IV. LE CENTRE DE LIMEIL-VALENTON (CEL-V)	6
V. LE DEPARTEMENT DE MATHEMATIQUES APPLIQUEES (D.A.M.)	7
VI. LE SERVICE S.I.E.	7
VII. PRESENTATION DU MATERIEL UTILISE	7
VII-1. Le DPS7-727 = LYRA	7
VII-2. Le système d'exploitation d'ordinateur BULL de la gamme DPS7 : GCOS7-LS	8
2ème Partie : DEVELOPPEMENT DE L'APPLICATION AVEC LE LOGICIEL MISTRAL	11
I. LE LOGICIEL MISTRAL	12
I-1. Fiche technique	12
I-2. Description de MISTRAL	12
I-3. Le marché de MISTRAL	14
II. ANALYSE ET SPECIFICATIONS AVEC LE CLIENT	15
II-1. Pourquoi vouloir créer une base de données?	15
II-2. Contenu de cette base	15
II-3. La base de données documentaire et ses acteurs	17
II-3-1. L'administrateur	17
II-3-2. Le documentaliste	17
II-3-3. L'utilisateur	18
III. CREATION DE LA BASE DE DONNEES	18
III-1. La PRécréation (PR)	19
III-1-1. Les CHamps (CH)	20
III-1-2. Les LExiques (LE)	23
a) <i>Création</i>	23
b) <i>Quelques explications supplémentaires</i>	24
III-1-3. DOmaines (DO)	26
a) <i>Définition des domaines</i>	26
b) <i>Création</i>	27
III-1-4. SEcret (SE)	27
III-1-5. Paramètres Divers (PD)	29
III-1-6. Champs Synonymes (CS)	30
III-1-7. Editions Cataloguées (EC)	30
a) <i>Création</i>	30

<i>b) Paramètres</i>	31
III-1-8. Relations (RL)	31
III-2. La Configuration (CO)	32
III-2-1. Tailles estimées des fichiers paramètres	33
III-2-2. Définition des différents fichiers	35
<i>a) Fichiers des Documents</i>	35
<i>b) Fichier des Lexiques</i>	35
<i>c) Fichier des Relations</i>	36
<i>d) Fichier Inverse des Descripteurs</i>	36
<i>e) Fichier Inverse des Synonymes</i>	37
<i>f) Fichier des Liens</i>	37
III-2-3. Organisation générale	38
III-3. La CRéation (CR)	39
III-4. STATistiques sur les informations contenues dans les Fichiers Permanents (ST)	42
III-5. Création des premiers formats de saisie	44
III-5-1. Principe de fonctionnement logue	44
III-5-2. Création des formats de saisie pour notre base	45
III-6. Création d'un fichier des mots vides pour les titres	47
III-6-1. Définition du fichier de mots vides	47
III--2. Création du fichier mot vide	48
III-6-3. Modification des séparateurs (SP)	49
IV. SAUVEGARDE DE LA BASE	51
V. MISE EN MARCHE ET ESSAI D'INTERROGATION	51
V-1. Recherche sur référence	52
V-2. Recherche par intervalle de valeur	53
V-3. Recherche sur fichier inverse	54
V-4. Recherche par auteur	56
3ème Partie : AUTRES POSSIBILITES DE MISTRAL	57
I. LA SAISIE	58
I-1. Les fonctionnalités du système	58
I-1-1. Création de documents	58
I-1-2. Modification - Mise à jour de documents	58
I-1-3. Suppression de documents	59
I-1-4. Visualisation des documents	59
I-2. Activation du module de Saisie/mise à jour	59
I-3. Les différents types de grilles nécessaires pour la création d'un document	60
I-4. Description détaillée des grilles	61
I-4-1. Grille 1 : choix des options générales de traitement	61
I-4-2. Grille 2 : traitement d'un groupe de champ 'court'	62
I-4-3. Grille 3 : traitement d'un champ 'long'	64
I-4-4. Grille 4 : fonction éditeur de texte	64
I-4-5. Grille 5 : choix des options en fin de traitement	64
I-5. Règles à savoir pour la saisie	66
II. COMMANDE DE MODIFICATION DES CHAMPS LONGS	66
II-1. Insertion de chaînes de caractères	69
II-2. Suppression de chaînes de caractères	70

II-3. Remplacement de chaînes de caractères	71
II-4. Compression de blancs	72
III. INTRODUCTION DE SYNONYMES DANS UN THESAURUS EXISTANT	73
III-1. Définition d'un thésaurus	73
III-2. La relation d'équivalence = la synonymie	74
III-3. Programme qui permet d'ajouter les synonymes au thésaurus existant	74
III-3-1. Organisation générale	75
III-3-2. Le programme FRTHES	75
III-3-3. Etape de vérification	77
a) Interrogation sur le thésaurus	77
b) Interrogation sur le lexique	78
CONCLUSION	80
ANNEXE 1 : ordre et signification des sigles des procédures	
ANNEXE 2 : liste MOTVID	

CONCEPTION D'UNE BASE DE DONNEES DOCUMENTAIRE AVEC LE LOGICIEL MISTRAL

Hélène BREUL

Stage effectué du 5 Juin au 28 Septembre 1990
au C.E.A., Commissariat à l'Energie Atomique, Département de
Mathématiques Appliquées, CEL-V, BP-27, 94195 Villeneuve-ST-
Georges, Tél.(1)45-95-65-22, sous la direction de Madame
Hourtoulle.

RESUME

La conception d'une base de données documentaire nécessite 4
étapes : étude de l'application (dossier de presse) avec
analyse avec le client, étude du logiciel de recherche
documentaire (MISTRAL), implantation de cette base sur
BULL/DPS7 et démarrage opérationnel.

DESCRIPTEURS : Base de données, MISTRAL, documentation,
informatique.

ABSTRACT

The conception of an information database necessitates 4
stages : study of application (press records) with analysis
with the client, study of information retrieval system
software (MISTRAL), implementation of this database on
BULL/DPS7 computer and execution.

KEYWORDS : database, MISTRAL, documentation, data processing.

Possibilité de diffusion

Confidentiel

REMERCIEMENTS

Je remercie les personnes qui m'ont permis d'effectuer ce stage de 4 mois en particulier, Madame Hourtoulle et Madame Inhauser qui se sont occupées de moi pendant toute la durée de ce stage ainsi que l'équipe du service S.I.E. qui, par leur aide et leur gentillesse, ont contribué à son bon déroulement. Je tiens aussi à remercier Madame Lelevier et Monsieur Drapeau qui travaillent au centre de Bruyère Le Châtel et qui ont bien voulu nous recevoir et nous donner les renseignements nécessaires à la création de la base.

INTRODUCTION

J'ai effectué un stage de 4 mois de juin à septembre 1990 au Centre d'Etude de Limeil-Valenton du CEA (Commissariat à l'Energie Atomique) au sein du service S.I.E. (Système Informatique et Exploitation) dépendant du Département de Mathématiques Appliquées.

Ce stage consiste en :

- la conception d'une base de données documentaire des dossiers de presse avec analyse et spécifications avec le client,
- implantation de cette base sur machine BULL/DPS7 avec le logiciel de recherche documentaire MISTRAL,
- recette avec le client.

PRESENTATION
DE
L'ENTREPRISE
ET
DES MOYENS INFORMATIQUES

PREMIERE PARTIE

I. LE COMMISSARIAT A L'ENERGIE ATOMIQUE (C.E.A)

Le CEA fut créé par le Général De Gaulle par ordonnance du 18 octobre 1945.

Il a comme mission de mettre en oeuvre l'utilisation de l'énergie atomique dans les domaines de la Science, de l'Industrie et de la Défense Nationale.

Le CEA est un établissement public de recherche et développement à vocation scientifique, technique et industrielle. Il exerce ses activités dans les domaines suivants : production de matières nucléaires, fabrication de réacteurs, recherche fondamentale, applications militaires, protection et sûreté nucléaire et enfin innovation et valorisation industrielles.

Le CEA est responsable sur les plans scientifique et technique des recherches et réalisations nucléaires et apporte à ce titre sa propre contribution à la conception des armements.

Il est constitué :

- d'un administrateur général qui assure la Direction Générale du CEA,
- d'un Haut-Commissaire qui assure les fonctions de conseiller scientifique et technique et exerce des responsabilités particulières dans le domaine de la sécurité, notamment en matière de protection et de sûreté nucléaire ainsi que dans celui de l'enseignement
- de Directions, à vocation générale ou scientifique. Les directions ont sous leur dépendance un certain nombre de centres destinés soit à la production, soit à la recherche.

II. LES CENTRES

Le Siège du CEA Civil et celui de la DAM, bien que distincts, sont situés dans le même bâtiment à Paris (XVème). Mais les centres sont volontairement répartis sur l'ensemble du territoire national.

Civils ou militaires, les Centres sont au nombre de 13 :

* Les centres civils :

- Cadarache
- Fontenay aux Roses
- Grenoble
- Saclay
- Vallée du Rhône : Pierrelatte et Marcoule

* Les centres de la DAM :

- Le Centre d'Etudes de Bruyères le Chatel (BIII)
- Le centre d'Etudes Scientifiques et Techniques d'Aquitaine (CESTA)
- Le centre d'Etudes de Limeil-Valenton (ÇEL-V)
- Le Centre d'Etudes du Ripault (CER)
- Le Centre d'Etudes de Valduc (CVA)
- Le Centre d'Etudes de Vaujours-Moronvilliers (CEV-M)

- Le Centre d'Essais de Villacoublay qui est couplé avec le CEP : Centre d'Expérimentation du Pacifique.

III. LA DAM (Direction des Applications Militaires)

La Direction des Applications Militaires est le principal artisan de la dissuasion nucléaire, c'est une branche très importante du CEA. Elle dispose d'un important potentiel scientifique et technique. Sa vocation première est de participer à la conception d'armes nucléaires, les plus efficaces, contribuant à la crédibilité de la force de dissuasion.

* Ses principaux domaines d'activité

- * *activités scientifiques* :
 - physique nucléaire
 - physique des lasers
 - calcul scientifique
- * *activités technologiques* :
 - matériaux de base
 - détonique
 - physico-chimie
- * *expérimentations* :
 - techniques générales de mesure
 - essais nucléaires
 - sûreté des expérimentations
- * *sécurité et protection* :
 - analyse des risques
 - prévisions
 - contrôles

IV- LE CENTRE D'ETUDES DE LIMEIL-VALENTON (CEL-V)

Le CEL-V, implanté dans le Val de Marne, est un établissement relevant de la DAM.

L'aspect "Défense" oblige à protéger fortement une large partie des activités du centre. Ces activités se répartissent équitablement entre la théorie, l'expérimentation et la simulation numérique sur ordinateurs :

- théorie : études scientifiques dans les nombreux domaines de la physique et la conception des charges nucléaires.
- expérimentation sur les plasmas et étude des interactions laser-matière. C'est en effet au CEL-V qu'a été implanté le laser Phebus, un des plus puissant laser, construit en collaboration avec des américains en 1985.

- informatique : un parc d'ordinateurs très puissant (CRAY-ONE, CRAY X-MP, des SUN3 et 4 et bientôt le CRAY Y-MP).

En 1970, un premier réseau scientifique DAM fût créé reliant les différents centres de la France en leur fournissant un support informatique.

Le personnel, environ 950 agents, dont près de la moitié sont des ingénieurs et des cadres.

V. DEPARTEMENT DE MATHEMATIQUES APPLIQUEES (DMA)

Ce Département de Mathématiques Appliquées a été constitué durant l'année 1984. Il est composé de différents services, dont le service S.I.E. dans lequel j'ai travaillé.

VI. LE SERVICE S.I.E.

Il a des responsabilités informatiques concernant l'exploitation des systèmes.

Il est chargé du bon fonctionnement des ordinateurs.

Il se subdivise en 3 sections chargées de :

- l'exploitation des ordinateurs,
- l'informatique de gestion et documentaire,
- réseaux et systèmes scientifiques.

La première section regroupe essentiellement les opérateurs qui sont chargés du fonctionnement des machines.

Le groupe d'informatique de gestion et documentaire travaille à l'amélioration des possibilités des ordinateurs et satisfait aux besoins des utilisateurs qui sont chargés des applications de gestion.

C'est dans cette branche de ce service S.I.E. que j'ai effectué mon stage.

Le personnel de ce groupe est chargé d'une part de la comptabilité des ordinateurs DPS7 BULL, celui de Limeil-Valenton et des autres centres de la DAM, d'autre part de l'informatique documentaire.

Enfin, la dernière section est constituée d'ingénieurs dont la fonction est l'exploitation des ressources des systèmes CRAY-CDC, ordinateurs tournés vers les travaux scientifiques.

VII. PRESENTATION DU MATERIEL

VII-1. Le DPS7-727 = LYRA.

DPS = Distributed Processing System.

Les DPS7 se mettent au service d'un grand nombre d'utilisateurs informaticiens et non informaticiens qui peuvent travailler simultanément dans différents modes : traitement par lots local ou à distance, interactif et transactionnel.

Ils couvrent une large gamme de puissances qui va de 1 à 35 fois environ.
Ils peuvent travailler en réseau.

caractéristiques :

- nombre de processeur : 2.
- mémoire principale : 12 MO.
- volume disque : 9 GO.

Le réseau est conçu sous forme de maillage entre les différents centres de la France. Ce système de maillage permet en cas de panne d'une branche du réseau, d'établir la communication par l'intermédiaire d'autres branches. (voir figure 1).

VII-2. Le système d'exploitation d'ordinateur BULL de la gamme DPS7 : GCOS7.LS

Le GCOS7-LS (Large System) permet la gestion des biprocesseurs comme le DPS7.727 et d'un grand nombre de terminaux. Il possède un noyau qui gère toutes les ressources de base du système.

GCOS offre un large éventail de langage de programmation : COBOL, BASIC, PASCAL, C, ...

La caractéristique la plus importante de GCOS7 est son adaptation aux différents types d'activité informatique, que ces activités soient sur le même système ou réparties sur plusieurs systèmes.

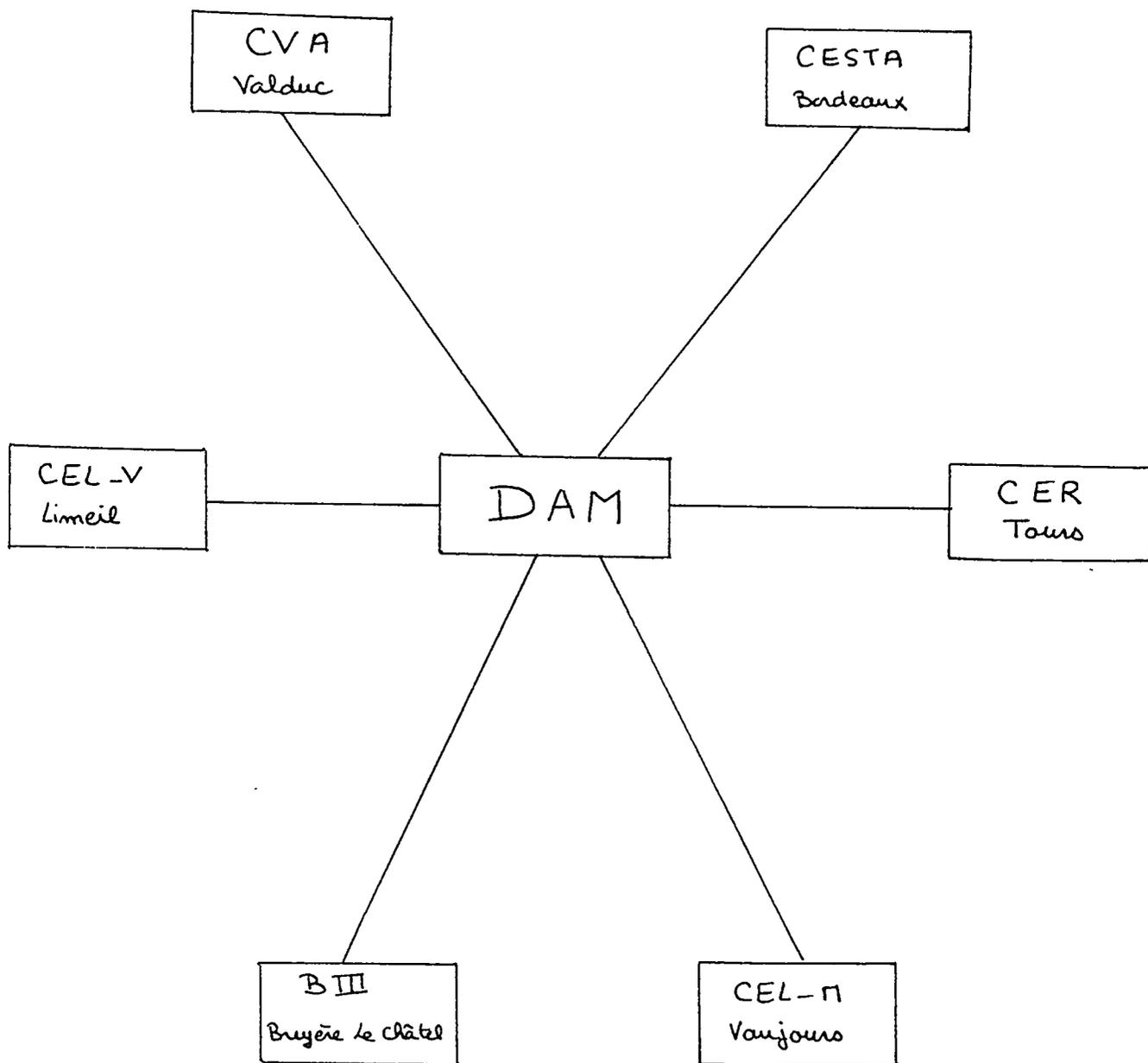
Il est articulé autour de 3 ensembles de produits adaptés aux 3 types d'activités que l'on rencontre en informatique :

- 'système de développement'. Certains produits logiciels permettent au service informatique de créer et de mettre à jour plus facilement des applications.

Le cycle de développement d'un programme étant une activité longue, impliquant des essais et des modifications avant d'obtenir un programme opérationnel, GCOS en réduit très sensiblement la durée en offrant des produits interactifs, tels que des compilateurs, des éditeurs de texte, des outils de mise au point de programme.

- 'système de production' qui comprend les outils nécessaires à la gestion et au traitement de grands volumes de données avec la meilleure sécurité possible.
 - * gestionnaire de bases de données (IDS II);
 - * moniteur transactionnel (TDS).
- 'système d'information' qui assure l'accès aisé des utilisateurs non informaticiens au DPS7 pour les aider à résoudre leurs problèmes d'informations.

FIGURE_1 : RESEAU DAM



Ce système est accompagné d'un logiciel de recherche documentaire (MISTRAL).

Il comprend donc des outils de :

- manipulation et interrogation de données (factuelles et textuelles);
- création d'applications informationnelles;
- visualisation graphique des données.

L'explication de MISTRAL est donnée dans la deuxième partie.

DEVELOPPEMENT
DE L'APPLICATION
AVEC LE LOGICIEL
MISTRAL

DEUXIEME PARTIE

I. LE LOGICIEL MISTRAL

MISTRAL = Mémorisation d'Information, Sélection, Traitement et Recherche Automatique par Logiciel.

MISTRAL est donc un logiciel de recherche documentaire qui est développé par BULL. Il constitue le coeur du système d'informations destiné à gérer les données bibliographiques, textuelles et factuelles de l'entreprise.

I-1. Fiche technique

- version du produit : V5.2-1A
- mode d'utilisation : IOF = "Interactive-Operating-Facility" (Ensemble de services interactifs en temps partagé facilitant les échanges entre un utilisateur et le système d'exploitation GCOS).

I-2. Description de MISTRAL

MISTRAL offre un ensemble de programmes permettant la mémorisation et la recherche d'informations textuelles ou factuelles. Système entièrement paramétrisé, il permet l'optimisation des traitements selon la nature des informations mémorisées.

Orienté recherche documentaire, il déborde le cadre habituel de celle-ci et autorise la gestion de type économique et statistique.

La grande capacité de stockage d'informations, sa gestion des lexiques et thésaurus, la puissance et la facilité d'utilisation de son langage de requêtes en font l'outil le plus adapté à la demande. Certaines fonctionnalités techniques de MISTRAL le mettent au top-niveau des produits de sa catégorie :

- ergonomie d'accès;
- intégration des minitels et des stations BULLQUESTAR 400 pour la saisie et la mise en forme;
- gestion de la confidentialité;
- ouverture sur les applications et les données de production.

Mistral réalise l'intégration du processus complet du traitement de l'information :

- production de l'information (saisie, indexation);
- organisation de l'information;
- services divers (services documentaires et services généraux de gestion des bases).

MISTRAL traite les informations pouvant être de 3 natures :

- les ensembles de mots-clés qui forment l'essentiel des bases bibliographiques. En effet, ces mots-clés servent à décrire le sujet du document : articles scientifiques, thèses, articles de presse, lois,

- réglements ou encore des objets : peinture, sculpture, photo etc...;
- les textes : textes "pleins" s'il s'agit par exemple du résumé d'une thèse, texte "intégral" s'il s'agit d'un article de loi par exemple;
 - les données numériques, dates et codes qui accompagnent fréquemment les textes (bases de données factuelles).

** Les fichiers*

Il existe des fichiers de documents, de synonymes, de lexiques et inverses. La décomposition en fichiers fait partie intégrante de MISTRAL.

** Souplesse et simplicité d'emploi de MISTRAL*

MISTRAL favorise l'accès des bases de données aux utilisateurs quelque soit leur niveau d'expérience. Il leur assure un service simple et rapide à l'information dont ils ont besoin. En effet, l'utilisateur "grand public" non familiarisé avec le langage d'interrogation et les bases qu'il interroge, dispose d'un menu d'aide qui peut l'assister à tout instant, ne serait-ce que pour lui donner la liste des lexiques d'interrogation.

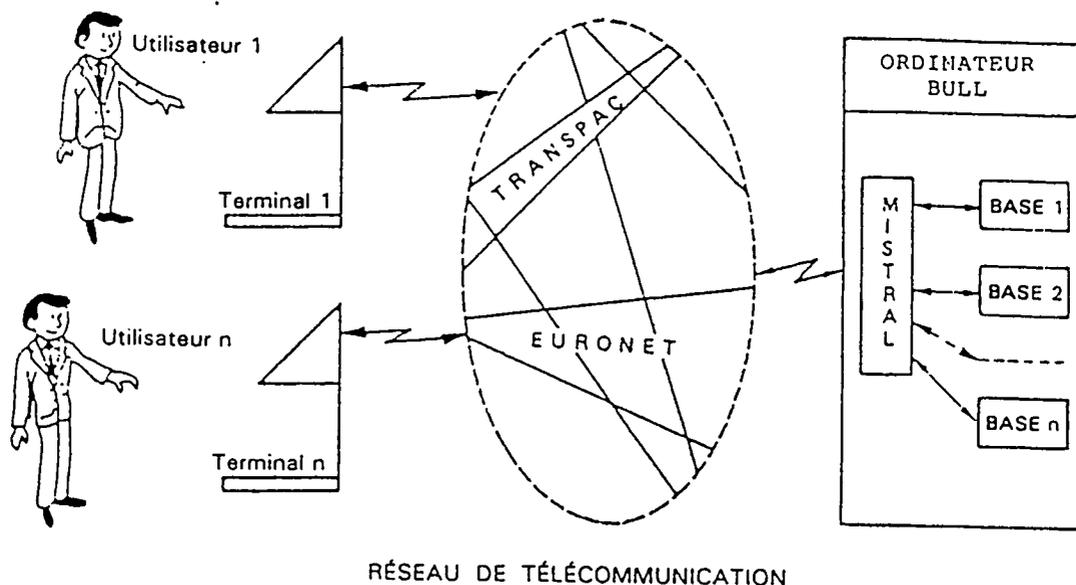
Le chercheur spécialisé, par contre, utilisera toute la puissance des outils de MISTRAL. Par exemple, il mettra en oeuvre le thésaurus afin d'obtenir plus rapidement l'information désirée.

** Travail dans sa propre langue*

MISTRAL permet de travailler dans plusieurs langues. Mais, le choix de l'utilisateur pour la langue ne se limite pas à la langue dans laquelle les documents sont stockés. En effet, MISTRAL accepte les demandes et affiche les messages dans la langue de son choix, ce qui facilite l'échange d'informations entre les bases situées dans les différents pays.

** MISTRAL et les réseaux d'informations*

MISTRAL fournit un accès rapide à l'information et ceci sans préjuger de l'endroit où elle se trouve. En effet, le logiciel des ordinateurs BULL permet aux utilisateurs situés aux endroits les plus divers d'interroger les bases de données grâce aux réseaux privés ou publics tels que TRANSPAC ou EURONET.



MISTRAL et les réseaux

I-3. Le marché de MISTRAL

MISTRAL est considéré dans de nombreux pays comme un outil de recherche documentaire d'une grande souplesse puisqu'il est utilisé aussi bien dans les banques que dans l'industrie pharmaceutique, l'agriculture, la papeterie... . Il se distingue donc par l'étendue pratiquement illimitée de ses applications et son caractère international.

** dans le monde*

MISTRAL est associé à un système de vidéodisque : une image de chaque sujet a été enregistrée sur un vidéodisque afin que l'utilisateur puisse véritablement voir apparaître une oeuvre d'art sur son écran pendant qu'il reçoit des informations sur l'origine et l'endroit où elle se trouve.

** dans la communauté internationale*

MISTRAL est utilisé par le CEE pour sa documentation interne mais aussi pour fournir des informations aux utilisateurs externes.

La CEE utilise une base de données appelée CELEX.

* *dans le traitement de l'information*

Les centres de Support Technique ainsi que Honeywell Informations Systems aux Etats-Unis utilisent une base de données par l'intermédiaire de MISTRAL.

II. ANALYSE ET SPECIFICATIONS AVEC LE CLIENT

Le but de ce stage consiste à la mise en place d'une base de données sur les dossiers de presse du CEA de Bruyère Le Châtel.

Comme toute démarche informatique, il est essentiel avant d'implémenter la base MISTRAL de bien connaître ses besoins et de définir ses objectifs.

Je suis donc allée avec Madame Hourtoule à Bruyère Le Châtel rencontrer Madame Lelevier et Monsieur Drapeau qui demandaient la mise en place de cette base de données avec MISTRAL.

II-1. Pourquoi vouloir créer une base de données?

Avant cette création de la base, le documentaliste faisait la saisie des articles de presse sur un logiciel MICRODOC mais il y a actuellement saturation. Cette saisie permettait l'édition d'un bulletin hebdomadaire de signalisation d'articles de presse. La documentation a donc décidé d'effectuer la saisie et l'édition de ces bulletins sur le logiciel DBASEIII.

L'informaticien du centre de Bruyère Le Châtel fera ensuite le transfert de ces notices de DBASEIII sur le DPS7 MISTRAL où sera créée la base de données.

Cette base de données sera donc utilisée pour :

- faire des recherches, interrogations;
- faire de la saisie;
- faire des mises à jour et des corrections;
- la gestion des lexiques.

II-2. Contenu de la base

Cette base de données contiendra des articles de presse (de caractère technique, politique, scientifique) ainsi que quelques ouvrages et rapports. Ces notices sont en français et en anglais sauf les résumés et les mots-clés qui sont toujours en français.

Cette base est sensée être utilisable pendant 10 ans environ.

Actuellement, les documentalistes saisissent 2000 documents par an. Cette saisie se fait depuis 4 ans; ce qui fait qu'il y a donc sur leur micro environ 8000 documents.

Cette base sera accessible à toute personne qui fait partie de la DAM. L'interrogation se fera par les documentalistes.

Le logiciel MISTRAL laissant une grande liberté de choix des structures (choix du nombre de champs, taille, intitulé, présentation, type de caractère qu'ils contiennent), on peut donc créer une application à ses propres besoins.

Avec les personnes responsables de l'application, on a défini les champs que contiendra cette base de données.

Champs	Abréviation	Nombre de lignes	type	lexique
Titre	TIT.	3	alphanumérique	oui
Auteurs	AUT.	2	alphanumérique	oui
Source	SOU.	2	alphanumérique	
Date	DAT.	8 car.	date	oui
Mots-clés	MCL.	3	alphanumérique	oui
Classement	CLA.	2	alphanumérique	
Disponible	DIS.	C,O,N	alphanumérique	
Intérêt	INT.	*,-	alphanumérique	
Résumé	RES.	14	alphanumérique	
Bulletin	BUL.	3 car.	alphanumérique	
Ouvrages	OUV.	O,N	alphanumérique	
Référence	REF.	p9999999	alphanumérique	

Les champs qui seront entrés dans la base doivent être de 4 lettres (ex. TIT.). Cette règle a été défini par le service. Les lexiques permettent de faire une interrogation sur le champ contenant le lexique.

Une explication des lexiques est donnée page

* Explication de quelques champs

SOU. C'est la source.
Si on a un ouvrage, on mettra ici le n° ISSN.
Pour un rapport, on met son numéro.

CLA. C'est le classement.
C'est un nombre correspondant aux différents chapitres où sont classés les documents. Il y a actuellement 23 codes de classement.

CHA. Ce champ permet de faire apparaître sur l'écran le libellé en clair du code de classement.

DIS. Ce champ comprend 3 caractères :
O = Oui
N = Non
C = Commandé

En effet, souvent un ouvrage commandé est annoncé dans le bulletin avant que celui-ci soit déjà reçu à la bibliothèque.

Il faudra donc dès la réception de la commande, modifier ce champ. Cette correction peut se faire directement sur la base et sur MISTRAL.

INT. On met "*" devant les articles qui sont jugés comme les plus intéressants sinon il y aura "-".

BUL. Ce champ correspond au numéro d'édition du bulletin. Ce numéro est séquentiel. Il y a environ 50 numéros par an.

OUV. Dans ce champs, on indique si on a un ouvrage ou un rapport.

REF. MISTRAL demande pour ce champ au minimum 8 caractères.

* Nous avons questionné les documentalistes au sujet de la présentation des notices à l'écran. En effet, on désirait savoir si elles voulaient voir les notices en majuscules ou en minuscules. Pour les documentalistes, cela ne leur fait rien car cette base va surtout servir à l'interrogation; l'édition des bulletins se faisant sur DBASEIII.

* Enfin, les documentalistes n'ont pas de choix préférentiel pour l'ordre d'apparition des champs. On est seulement obligé par MISTRAL de mettre la référence en première ligne à l'écran.

II-3. La base de données documentaire et ses acteurs

II-3-1. L'administrateur

Généralement, cette personne est un informaticien. Il a à sa charge la maintenance de la base, en gros les sauvegardes, la configuration des bases et tout ce qui touche au système hôte. C'est lui qui crée la base de données. Il définit les environnements de travail des utilisateurs (identification, bases accessibles, droits d'accès, procédures autorisées...), le vocabulaire de travail, les formats de saisie et les formulaires d'interrogation.

Il doit être à même de conseiller sur les différents types de matériels que l'on peut connecter et sur les procédures à adopter.

L'administrateur, c'est le service support/maintenance de la base de données.

II-3-2. Le documentaliste

Le documentaliste est la personne qui alimente la base de données en documents. Il classe les documents, renseigne les divers champs utiles à chaque document.

Il est responsable de l'évolution de la base de données; c'est lui qui la fait évoluer par les créations, mise à jour et suppressions de documents.

Le documentaliste, c'est le fournisseur de la base.

II-3-3. L'utilisateur

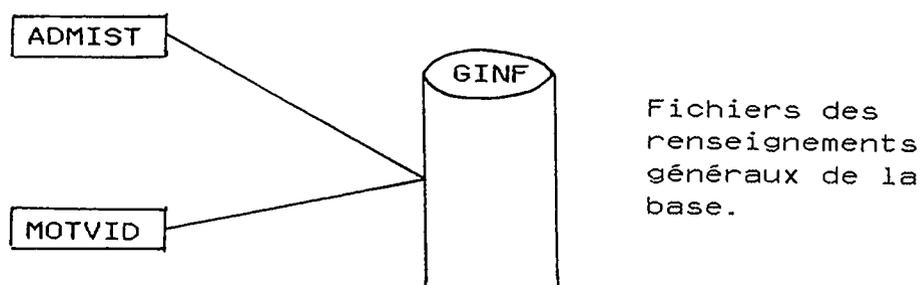
Son activité consiste à rechercher des documents et surtout à les trouver.

L'utilisateur, c'est le client de la base de données.

III. CREATION DE LA BASE DE DONNEES

Pour créer une base de données, il faut utiliser sur MISTRAL :

- le programme ADMIST (programme d'administration des bases) qui permet à l'administrateur de déclarer les différents formats de saisie, les paramètres, les champs, domaines et habilitation des usagers.
- le chaîne MOTVID (traitement par lots) qui permet la constitution du dictionnaire de mots vides.



ADMIST est donc un outil interactif conversationnel de configuration et de suivi d'exploitation d'un stock documentaire mis à la disposition de l'administrateur.

On distingue deux niveaux de traitements :

- AD = ADministrateur au niveau d'une base documentaire.
- GE = GEstion au niveau de l'ensemble des bases d'un stock documentaire.

J'utilise donc le niveau AD qui comprend différentes parties :

- définition des paramètres formels et structurels d'une base.
- création des fichiers permanents propre à une base.
- mise à jour des paramètres de création d'une base.
- lancement de la chaîne de création des mots vides associés à une base.
- lancement de l'édition des paramètres de la base.

ADMIST offre à l'utilisateur 7 grandes fonctions :

- la PRécréation (PR) : cette fonction est la seule possible à l'initialisation du fichier des renseignements généraux (GINF);
- la COnfiguration (CO);
- la mise à jour avant création de fichiers (MA);
- la CRéation des fichiers permanents (CR);
- la VISualisation (VI) suivi de l'étape voulue séparée par un blanc;
- l'édition de STatistiques (ST);
- la Mise à Jour après création des fichiers (MJ).

Appel du programme ADMIST

Dès le lancement du programme ADMIST, un dialogue s'échange avec l'administrateur des bases. L'administrateur choisit le niveau "Administration" d'une base (AD) et indique le numéro de la base concernée. Le système propose alors la liste totale de toutes les procédures possibles.

Le lancement du programme ADMIST est montré ci-dessous.

```
S: admist
B U L L MISTRAL V5.2-1A
WORK LANGUAGE?
FR EN
?fr
niveau de traitement?
?ad
numero de la base (entre 1 et 25)?
+++

?13
nom de la base(7caracteres oblq)°,mot de passe(<=8caracteres)§
?base@13
support du fichier des R.G.:media,devclass?
?mist1,ms/d50@
mot de protection :

AXXXXXXXXXXXXXX

fonction ?
???
```

III-1. La PRécréation (PR)

Cette fonction permet par l'exécution d'étapes successives d'initialiser et de modifier tous les paramètres qualitatifs de MISTRAL (vocabulaire, séparateurs, relations, ...) et d'introduire des paramètres de nature documentaire (nombre de documents par domaine, nombre de termes d'un lexique ...) qui serviront lors de la configuration.

Etapas de la précréation

- MP : définition ou modification d'un Mot de Passe pour la protection de la gestion de la base en traitement par lot.
- RL : définition ou modification des ReLations entre documents ou entre descripteurs de la base.
- CH : définition ou modification des noms ou caractéristiques des CHamps.
- EC : définition ou modification des Catalogues de champs à Editer
- CS : définition ou modification des Champs Synonymes.
- LE : définition ou modification de la liste des champs auxquels doivent correspondre les descripteurs composant un LExique.
- DO : définition ou modification des DOmaines (type périodique ou non, nombre de documents en création ou en mise à jour, mot de passe du domaine ...)
- PD : définition ou modification des Paramètres Divers.
- SE : définition ou modification du niveau de SEcret.

III-1-1. Les CHamps (CH)

Le champ référence doit être le premier créé de tous, il n'est constitué que par son indicatif et son libellé.

* création du champ référence

? indch, libch (indicatif et libellé entrés)

avec

indch : indicatif du champ de 1 à 8 caractères.

libch : libellé documentaire du champ de 1 à 16 caractères

* création des autres champs

? indchp, lidchp, secchp, $\left. \begin{array}{l} \text{oblg} \\ \text{opt1} \end{array} \right\}$, $\left[\begin{array}{l} \text{alph} \left[\begin{array}{l} \text{,msep} \\ \text{,arlt} \end{array} \right] \text{,lexu} \\ \text{num} \left\{ \begin{array}{l} \text{,arlt} \\ \text{,msep} \end{array} \right\} \text{,lexu} \left\{ \begin{array}{l} \text{,lexu} \\ \text{,lexu} \end{array} \right\} \end{array} \right]$

Paramètres :

- indchp : indicatif de champ (1 à 8 caractères).
- libchp : libellé documentaire du champ (1 à 16 caractères).
- secch : niveau de secret, nombre compris entre 0 et 49.
- alph : champ alphanumérique.
- num : champ numérique.
- date : champ type date.
- oblg : champ obligatoire.
- optl : champ optionnel.
- msep : champ multiséparateur. Les valeurs de ce champ sont alors les mots composant le texte et la table des séparateurs contient le caractère blanc et tous les caractères de ponctuation. (on a utilisé ce champ pour les titres, ce qui fait que tous les mots du titre sont indexés comme mots-clés)
- lexu : champs à incorporation automatique des mots inconnus (facultatif)
- arlt : traitement automatique des relations sur groupes de mots. (association d'un groupe de mots à ces composants)(facultatif). Il permet le pointage. On l'utilise pour les mots-clés. Si dans un lexique MCL on a le mot '*alliage' cela veut dire qu'il y a d'autres mots composés qui contiennent le mot alliage
Ex. Alliage électronique... .

L'option 'arlt' de traitement automatique des relations est interdite pour les champs multiséparateurs.

Le nombre maximum de champs introduits (y compris les champs synonymes) est de 250.

L'indicatif, le libellé et le niveau de secret doivent être entrés les premiers dans cet ordre, les autres paramètres peuvent l'être dans un ordre quelconque.

L'accès aux champs d'un document se fait par leurs indicatifs.

La création des champs et un tableau récapitulatif sont donnés ci-dessous.

creation/modification d'un champ :
?ref,reference

creation/modification d'un champ :
?tit.,titre,@,optl,alph,msep,lexu

creation/modification d'un champ :
?aut.,auteurs,@,optl,alph,lexu

creation/modification d'un champ :
?sou.,source,@,optl,alph

creation/modification d'un champ :
 ?dat.,date,@,opt1,date,lexu

creation/modification d'un champ :
 ?mcl.,mots-cles,@,opt1,arlt,lexu
 nature du champ incorrecte(ALPH/NUM/DATE/NOTX)

creation/modification d'un champ :
 ?mcl.,mots-cles,@,opt1,alph,arlt,lexu

creation/modification d'un champ :
 ?cla.,classement.@,opt1,alph

creation/modification d'un champ :
 ?dis.,disponible,@,opt1,alph,

creation/modification d'un champ :
 ?int.,interet,@,opt1,alph

creation/modification d'un champ :
 ?bul.,bulletin.@,opt1,alph

creation/modification d'un champ :
 +++

?ouv.,ouvrage,@,opt1,alph

creation/modification d'un champ :
 ?cha.,chapitre,@,opt1,alph

creation/modification d'un champ :
 ?res.,resume.@,opt1,alph

creation/modification d'un champ :
 ?

?vi ch

champs

```

*****
: indic :cod:lex: libelle :ns:nat.:obg:indx:lxu:arlt:s:
*****
:AUT. : 3: :AUTEURS : 0:ALPH: : :OUI: : :
:BUL. : 10: :BULLETIN : 0:ALPH: : : : : :
:CHA. : 12: :CHAPITRE : 0:ALPH: : : : : :
:CLA. : 7: :CLASSEMENT : 0:ALPH: : :OUI: : :
:DAT. : 5: :DATE : 0:DATE: : :OUI: : :
:DIS. : 8: :DISPONIBLE : 0:ALPH: : : : : :
:INT. : 9: :INTERET : 0:ALPH: : : : : :
:MCL. : 6: :MOTS-CLES : 0:ALPH: : :OUI:OUI: : :
:OUV. : 11: :OUVRAGE : 0:ALPH: : : : : :
:REF : 1: :REFERENCE : : : : : : : :
:RES. : 13: :RESUME : 0:ALPH: : : : : :
:SOU. : 4: :SOURCE : 0:ALPH: : : : : :
:TIT. : 2: :TITRE : 0:ALPH: :MSEP:OUI: : :
*****
  
```

II-1-2. les Lexiques (LE)

Un lexique est un dictionnaire alphabétique des termes d'indexation des documents.

Un lexique est associé à un ou plusieurs champs indexés.

On en déduit qu'à une base documentaire peuvent être associés plusieurs lexiques : la liaison entre le vocabulaire permettant de poser des questions et celui des documents est réalisée par les lexiques et thésaurus (ensemble des descripteurs d'un champ (lexiques monochamps ou multisens) ou de plusieurs champs (lexique monosens) d'un lexique d'une base documentaire. Ces descripteurs sont reliés entre eux par des relations sémantiques).

a) création

La définition d'un lexique se fait en trois stades différents:

- introduction des champs associés au lexique,
- nature monosens ou multisens,
- paramètres du lexique.

* introduction des champs

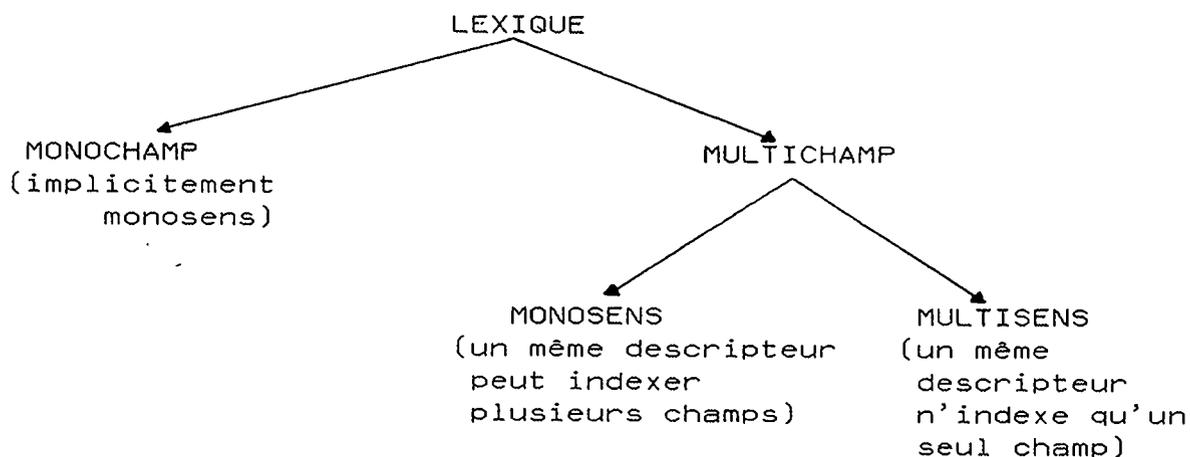
Si le lexique ne comprend qu'un seul champ, il est dit MONOCHAMP, sinon il est dit MULTICHAMP.

Les indicatifs des champs associés au lexique sont entrés séparés par des virgules.

* nature monosens ou multisens

Si tout descripteur du lexique peut avoir pour origine plusieurs champs, le lexique est dit MULTISENS, sinon il est dit MONOSENS.

Ainsi, les différents types de lexiques sont résumés dans le schéma général suivant :



Dans un même lexique, on ne peut pas mélanger des champs TEXT avec des champs non TEXT.

Un lexique des champs TEXT est forcément multisens ou monochamp.

* paramètres

Dans la précréation, des paramètres documentaires sont demandés en vue de la configuration :

- nombre de descripteurs à la création du lexique;
- nombre de descripteurs introduits en mise à jour;
- longueur moyenne d'un descripteur.

En mise à jour et après création, la création de nouveaux lexiques n'est possible qu'avec de nouveaux champs et dans la mesure où il reste une place suffisante dans les fichiers correspondants.

Si aucune place n'est disponible dans le fichier des Lexiques ou dans le fichier des Liens, la création est abandonnée.

Le nombre maximum de lexiques pouvant être introduits est 50.

Nous avons défini 5 lexiques qui sont : TIT., AUT., MCL., DAT., CLA. . Tous les champs associés au lexique sont visualisés par ordre croissant de leur code.

b) quelques explications supplémentaires sur les lexiques créés

* Champ : DAT.

entree des champs

?dat.

lexique monosens(O/N)?

?o

nombre de descripteurs a la creation

?365

nombre de descripteurs introduits lors des differentes maj

?365@

longueur moyenne d'un descripteur

?8

J'ai entré dans le 'nombre de descripteurs à la création' 365 ce qui correspond au nombre de dates possibles entrées dans la base.

On considère que cette base sera utilisée pendant 10 ans ce qui donne un 'nombre de descripteurs introduits lors des différentes mise à jour' de 3650.

La date est entrée comme suit : AA/MM/JJ ce qui correspond à 8 caractères.

* Champ : AUT.

entree des champs

?aut.

lexique monosens(O/N)?

?o

nombre de descripteurs a la creation

?800

nombre de descripteurs introduits lors des differentes maj

?8000

longueur moyenne d'un descripteur

?20

* Champ : MCL.

entree des champs

?mcl.

lexique monosens(O/N)?

?o

nombre de descripteurs a la creation

?500

nombre de descripteurs introduits lors des differentes maj

?1000

longueur moyenne d'un descripteur

?40

Le 'nombre de descripteurs introduits lors des différentes mise à jour' n'est que de 500 car le sujet qui est sur les Applications Militaires du Nucléaire est très pointu et précis ce qui restreint le nombre de descripteurs nouveaux à entrer.

* Champ : CLA.

entree des champs

?cla.

lexique monosens(O/N)?

?o

nombre de descripteurs a la creation

?30

nombre de descripteurs introduits lors des differentes maj

?10

longueur moyenne d'un descripteur

?2

Actuellement, ce classement comprend 23 parties et une annexe. Il comporte donc 1 à 5 chiffres (I-XXIII). Ce qui donne une longueur moyenne de 2.

* Champ : TIT.

entree des champs
+++

?tit.
lexique monosens(O/N)?
?o
nombre de descripteurs a la creation
?10000
nombre de descripteurs introduits lors des differentes maj
?10000
longueur moyenne d'un descripteur
?10

Les différents lexiques créés sont récapitulés dans le tableau suivant :

reprendre,visualiser ou prochaine etape (RE,VI nom1,nom2)?
?vi le

lexiques

```

*****
: indicatif : code : monosens : tail. lx(bl) : tail. ls(bl) :
*****
: REF      : 1 : OUI  :           :           :
: TIT.     : 2 : OUI  :           :           :
: AUT.     : 3 : OUI  :           :           :
: DAT.     : 5 : OUI  :           :           :
: MCL.     : 6 : OUI  :           :           :
: CLA:     : 7 : OUI  :           :           :
*****

```

III-1-3. DOmaines (DO)

a) définition des domaines

La structure principale du fichier des Documents est le domaine. Un domaine est la partition d'une base regroupant des documents par sélection de champs spécifiques de niveau de secret.

Le nombre de domaines possibles est limité à 50.
L'administrateur peut ne pas partager son domaine.

b) création

La création d'un domaine se fait en 2 stades différents :
 - introduction des caractéristiques du domaine,
 - introduction des paramètres.

* introduction des caractéristiques du domaine
 Sur l'écran on voit apparaître :

création/modification d'un domaine
 ? indem, mp : motpas

avec

indem : indicatif du domaine de 1 à 8 caractères.
 motpas : mot de passe de 1 à 8 caractères.

* introduction des paramètres

Les paramètres documentaires sont demandés en vue de la configuration :

- nombre de documents introduits à la création du domaine,
- nombre de documents introduits en mise à jour,
- longueur moyenne d'un document (en nombre de caractères),
- longueur moyenne des documents d'un champ texte (en nombre de caractères).

La création du domaine pour la base est donnée ci-dessous :

add

creation/modification d'un domaine :
 ?presse,ensor

-----!

parametre de longueur erronee(8cars max) ou manquant

?presse,mp:ensor

nombre de documents introduits a la creation

?2000

nombre de documents introduits lors des differentes maj

?10000

longueur moyenne des documents(en cars)

?500

III-1-4...SEcnet..(SE).

Cette partie sert à la définition des clés de secret et de leur niveau.

MISTRAL gère la confidentialité du contenu d'une base (documents) et du contenu des documents (champs) en vérifiant, dans la clé d'habilitation associée à son utilisateur,

l'existence d'une valeur correspondant au niveau de secret attaché aux documents ou aux champs accédés.

120 niveaux de secret sont possibles (de 0 à 119) :

- 0 = accès libre (pas de protection).
- 25 = accès strictement réservé aux utilisateurs dont la clé d'habilitation comprend la valeur 25.
- 119 = accès strictement réservé aux utilisateurs dont la clé d'habilitation comprend la valeur 119

NB : la clé d'habilitation exprime le droit d'accès :

- aux documents;
- aux champs des documents;
- aux fonctions du bibliothécaire;
- aux fonctions de la saisie.

* habilitation pour l'utilisation du bibliothécaire

A une clé peut correspondre 3 niveaux d'habilitation pour accéder au bibliothécaire :

- accès en lecture seulement (C.R.);
- accès en écriture, en création, sans modification (1);
- accès en écriture, en création, modification, suppression (2).

* habilitation pour l'utilisation de la saisie interactive

A une clé peut correspondre 3 niveaux d'utilisation de la saisie :

- saisie interdite (C.R.)
- saisie de documents, sans modification ou suppression de documents existants = création (1)
- saisie de documents, modification ou suppression de documents existants = MAJ (2)

La définition de la clé de secret pour la base est donnée ci dessous :

?2

creation/modification d'une cle de secret:

?brissa:0/0

BRISSA : 0/0

acces au bibliothec.: 1=creat./2=maj/C.R.=lect.seulmt (1/2/CR)?

?2

acces a la saisie: 1=creation / 2=maj / C.R.=interdit (1/2/CR)?

?2

III-1-5. Paramètres Divers (PD)

Liste des paramètres divers :

- introduction des documents par référence croissante (oui/non).
- introduction de la synonymie : pour adapter la taille du fichier inverse des synonymes (FSY) qui par défaut est seulement de 1 Unité Allocation (U.A.). (1 U.A.= 120 blocs de 1024 octets).
- utilisation des relations entre documents pour l'ajustement des fichiers des relations.
- longueur maximale de la référence (en nombre de caractères). La longueur maximale est un nombre n tel que $8 \leq n \leq 34$.
- format de la date du document. On a le choix entre JJ/MM/AAAA et AAAA/MM/JJ.
- longueur maximale de la partie entière d'un nombre en champ numérique.
- longueur maximale de la partie décimale d'un nombre en champ numérique.
- longueur maximale d'un document (en nombre de caractères).
- longueur maximale d'un descripteur (en nombre de caractères). Elle ne peut dépasser 245 caractères
- troncature des mots-clés trop long.
- nombre de colonnes d'édition = largeur de la ligne. Le nombre n de colonne d'édition est tel que $60 \leq n \leq 132$.

Les paramètres que l'on a entré pour notre base sont :

reprendre, visualiser ou prochaine etape (RE,VI nom1,nom2)?
?vi pd

parametres divers

introduction des documents par references croissantes(O/N)?....	0
utilisation de la synonymie(O/N)?.....	0
utilisation des relations entre documents(O/N)?.....	N
longueur maximale de la reference($8 \leq n \leq 34$)?	8
format de la date dans le document(jj/mm/aaaa=1;aaaa/mm/jj=2)?	2
longueur max de la partie entiere d'un nb en champ numerique?	5
longueur max de la partie decimale d'un nb en champ numerique?	2
longueur max d'un document?	2000
longueur max d'un descripteur(≤ 245)?	100
troncature des mots-cles trop longs(O/N)?.....	0
nombre de colonnes d'edition($60 \leq n \leq 132$)?	80

III-1-6. Champs Synonymes (CS)

La création d'un champ synonyme est la suivante :

? indch1, indch2

paramètres :

- indch1 : indicatif d'un champ déjà introduit (1 à 8 caractères).
- indch2 : indicatif synonyme du premier (1 à 8 caractères); cet indicatif ne doit pas déjà exister.

L'administrateur doit entrer trois champs synonymes de nom TXTCHPIMP, LEXIMP, THCHPIMP. Ces champs seront reconnus par MISTRAL respectivement comme champ implicite pour la recherche sur texte, lexique implicite et thésaurus implicite pour un lexique multisens.

(CS)

creation d'un champ synonyme
?mcl.,leximp

creation d'un champ synonyme
?mcl.,thchimp

creation d'un champ synonyme
?mcl.,thchpimp

creation d'un champ synonyme
?tit.,txchpimp

III-1-7. Editions Cataloguées (EC)

Une édition cataloguée est une liste de l'ensemble des champs dont l'édition peut être demandée en même temps. Un nom est affecté par l'administrateur à chaque édition. Ce nom sera rappelé par le gestionnaire des bases lors de l'initialisation d'un environnement utilisateur pour préciser l'édition implicite en début de session MISTRAL.

Au moins une édition cataloguée doit être introduite avant la création des fichiers d'une base.

a) création

message visualisé à l'écran :

création/modification d'une édition cataloguée

? edc, chp1, chp2, ...

b) paramètres

- edc : indicatif de l'édition cataloguée de 1 à 4 caractères.
- chp : indicatif de champ déjà introduit (1 à 8 caractères).

Le nombre d'édition cataloguée est limitée par la formule suivante :

$$NB = (1024 - \text{nombre total de champs créés dans les éditions})/6$$

Pour la base que je crée, on a décidé de donner 2 formats possibles de visualisation des documents recherchés : soit le format EC1, soit EC2.

Le contenu de ces éditions cataloguées est montrée ci-dessous:

reprendre,visualiser ou prochaine etape (RE,VI nom1,nom2)?
 ?ec

creation/modification d'une edition cataloguee
 ?ec1,ref,tit.,aut.,sou.,dat.,mcl.

creation/modification d'une edition cataloguee
 ?ec2,ref,tit.,aut.,sou.,dat.,mcl.,cla.,cha.,dis.,int.,res.,bul.,ouv.

reprendre,visualiser ou prochaine etape (RE,VI nom1,nom2)?
 ?vi ec

editions cataloguees

EC1

REF	TIT.	AUT.	SOU.	DAT.	MCL.
-----	------	------	------	------	------

EC2

REF	TIT.	AUT.	SOU.	DAT.	MCL.	CLA.
CHA.	DIS.	INT.	RES.	BUL.	OUV.	

111-1-8. ReLations (RL)

13 opérateurs de relation sont applicables aux relations entre descripteurs et entre documents d'une même base. MISTRAL privilégie les relations réciproques de hiérarchies (spécifiques, génériques) et la relation d'équivalence entre descripteurs (synonymie). Ces deux types de relation sont initialisés par défaut.

création

L'utilisateur entre ses relations de la façon suivante :

? op, {trm : sig1, lib1}

avec

- op : opérateur de 2 caractères au maximum.
- sig1 : sigle de 1 à 4 caractères de la relation entre descripteurs.
- lib1 : libellé de 32 caractères au maximum de la relation entre descripteurs (le caractère blanc est admis à l'intérieur du libellé).
- trm : qualifie une relation entre descripteurs.

L'utilisateur peut entrer jusqu'à 10 relations propres. Après la création des fichiers de la base, une nouvelle relation ne pourra être introduite que si le fichier des Relations est vide. Pour les relations privilégiées, une valeur par défaut est proposée. Dans le champ mots-clés (MCL.) de notre base, nous avons utilisé un champ à relation automatique qui est le paramètre ARLT. Cela veut dire que l'on a un traitement automatique des relations sur des groupes de mots (association d'un groupe de mots à ses composants). Par exemple, le mot *alliage* est associé avec le mot *alliage électronique*. Or, pour les champs à relation automatique, une relation de groupe de mots doit être entrée par l'administrateur. J'ai donc entré la relation ARLT pour le champ MCL..

7-1

relations privilégiées existantes:

>,TRM:BRDT,GENERIQUES	,DOC:PERE,PERE
<,TRM:NRWT,SPECIFIQUES	,DOC:FILS,FILS
=,TRM:SYN ,EQUIVALENTS GRAMMATICAUX	

creation/modification d'une relation

?:,trm:arlt

manque de données

?:,trm:arlt,groupe de mots

III-2. la COntfiguration (CO)

Cette fonction est lancée après la précréation. Elle assure la transformation des paramètres documentaires introduits lors de la précréation, en paramètres MISTRAL.

Le but de cette fonction est de fournir à l'administrateur une estimation de la place disque nécessaire pour installer la base en tenant compte des paramètres donnés pendant la phase de précréation (PR).

Pour que cette fonction puisse être exécutée, quatre étapes doivent obligatoirement avoir été lancées dans la précréation:

- LE : existence d'un lexique au moins.
- DO : existence d'un domaine au moins.

- PD : initialisation des paramètres divers.
- EC : existence d'une liste de champs à éditer (Edition Cataloguée).

La figure 2 montre l'initialisation des différents fichiers.

III-2-1. Tailles estimées des fichiers permanents

Les tailles des fichiers sont exprimées en blocs (1 bloc = 1024 octets).

Les tailles des fichiers débordement sont multiples de 6 blocs.

Les tailles estimées des fichiers des Documents principal et débordement sont les sommes des tailles des sous-fichiers correspondant aux différents domaines.

Les tailles estimées des fichiers des Lexiques et des Liens sont les sommes des tailles des sous-fichiers correspondant aux différents lexiques.

Les tailles des fichiers DOCM, LEXM et LINK sont des tailles minimales qui ne peuvent qu'être augmentées.

Ces tailles sont montrées sur le tableau :



taille estimée des fichiers

* fichier *taille(bls)*

: DOCM : 6840 :

: DOCO : 6 :

: LEXM : 1800 :

: LEXO : 354 :

: LINK : 1200 :

: RLTN : 720 :

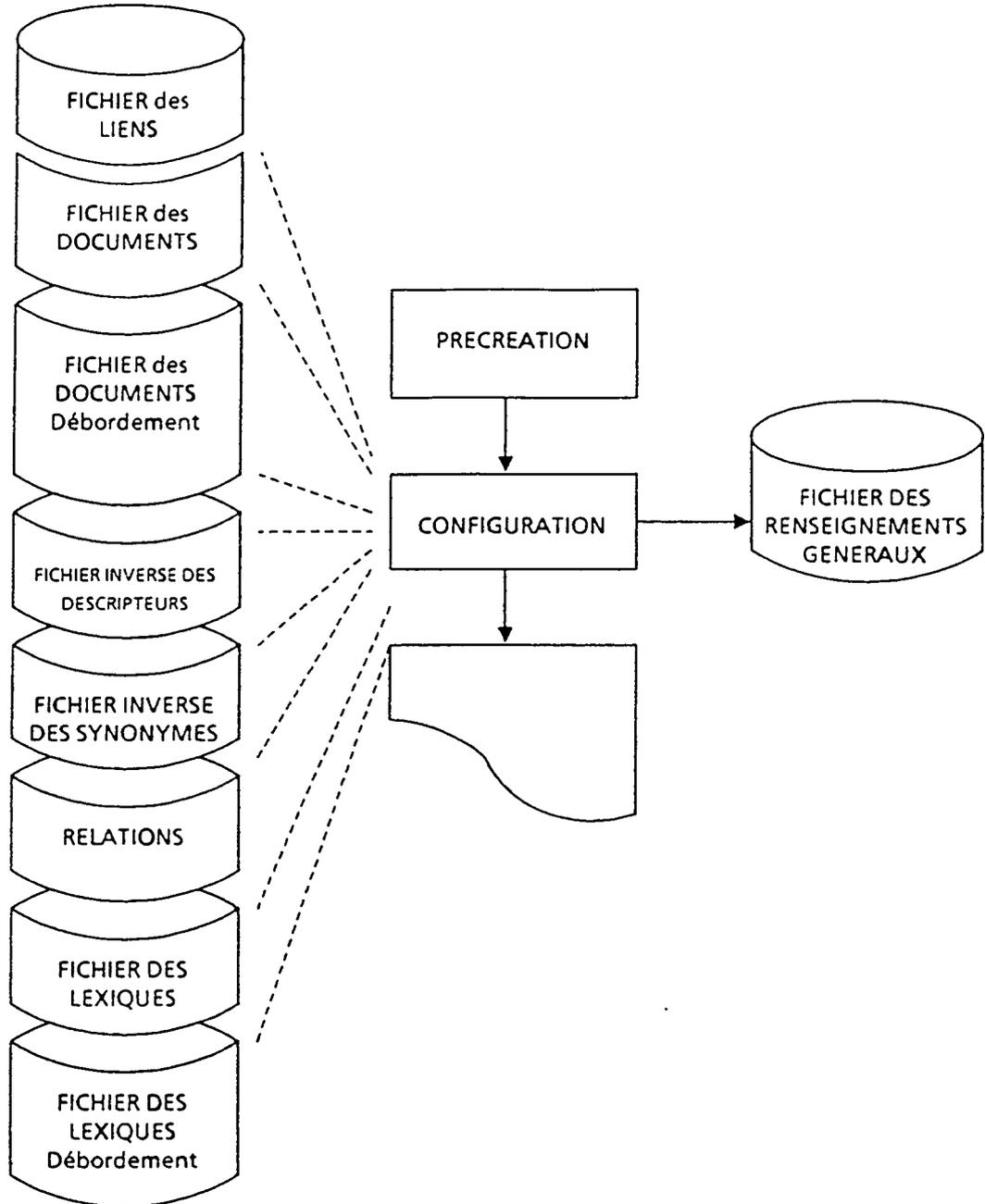
: FIDE : 6840 :

: FISY : 720 :

* développement des sigles utilisés :

- DOCM : fichier des Documents principal constituant la base.
- DOCO : fichier des Documents débordement.
- LEXM : fichier des Lexiques principal relatif à une base.
- LEXO : fichier des Lexiques débordement.
- LINK : fichier des Liens inter-fichiers d'une base.
- RLTN : fichier des Relations entre documents et/ou descripteurs d'une base.

FIGURE 2



INITIALISATION DES FICHIERS

- FIDE : fichier Inverse des Descripteurs d'une base.
- FISY : fichier Inverse des Synonymes des descripteurs d'une base.

III-2-2. définition des différents fichiers

a) Fichiers des Documents

Il contient les documents entrés dans la base documentaire. En pratique, il est composé de 2 fichiers :

- 1) le fichier principal des Documents constitué d'articles de longueur fixe : la cellule (une cellule par document). Ce fichier est subdivisé en domaines.

Chaque cellule est constituée de 2 parties :

- une zone en-tête comportant divers renseignements sur la structure du document.
- le texte du document (la totalité ou seulement le début de ce texte si ce dernier a une longueur supérieure à la partie de la cellule disponible à la suite de la zone en-tête).

- 2) le fichier débordement des documents qui contient la suite des documents.

Il n'y a aucun découpage du fichier débordement en domaines.

Un bloc du fichier débordement est découpé en 32 segments de longueur fixe. Une suite de documents est constituée, dans un bloc de débordement (1024 octets), par un ou plusieurs groupes de segments de 32 octets chaînés entre eux. La suite relative à un même document peut s'étendre sur n blocs consécutifs mais dans ce cas, elle est découpée en n groupes de cellules.

Donc, chaque article se présente sous la forme d'une cellule.

b) Fichier des Lexiques

Le fichier Lexique ne constitue qu'une partie de thésaurus proprement dit. Ce dernier regroupe les fichiers Lexique, Liens et Relations.

Le fichier Lexique est donc le fichier qui contient tous les lexiques relatifs à une base documentaire.

Il donne une liste des descripteurs avec pour chacun de ces descripteurs le numéro de code qui lui est associé. Ce numéro est le numéro d'ordre à la création des articles correspondants dans le fichier Liens. C'est sous ce numéro que le descripteur est représenté dans les fichiers Liens, Relations et les fichiers Inverses qui permettent de remonter au fichier des Documents.

Le fichier Lexique se décompose en 2 fichiers :

- Lexique principal lui-même subdivisé en Lexiques (chaque lexique correspondant à un champ ou à un groupe de champs associés au fichier inverse).

- Lexique débordement (extension du fichier principal).

Par ailleurs, sont stockés dans une zone réservée des documents ainsi que les pointeurs vers le fichier des Documents.

Structure générale du lexique

La dimension d'un lexique est un nombre entier d'Unités d'Allocations (U.A. = 120 blocs).

Chaque Unité d'Allocation comprend une partie principale et une partie débordement. Mais, si celle-ci est saturée, les articles sont alors écrits dans le fichier Lexique débordement général, commun à tous les lexiques.

Chaque lexique est organisé en index qui contiennent les descripteurs ainsi que divers renseignements statistiques.

c) Fichier des Relations

Il contient les relations définies par l'utilisateur entre les mots des lexiques, ainsi qu'entre les documents du fichier des Documents.

* relations entre descripteurs

Un article du fichier Relations regroupe dans ce cas les relations de hiérarchie et de synonymie associées à un descripteur sous forme de listes de spécifiques, de génériques, de synonymes de ce mot et/ou relations définies par l'utilisateur.

Si le descripteur appartient à une classe de synonymes, l'article relation est commun à tous les descripteurs de la classe.

* relations entre documents

Un article du fichier Relations regroupe alors les relations de hiérarchie associées à un document sous forme de listes spécifiques, de génériques de ce document et/ou de relations définies par l'utilisateur.

Pour la base que j'ai créée, j'ai utilisé la relation groupe de mots

d) Fichier Inverse des Descripteurs

Il contient, par descripteur, la liste des documents indexés. Le fichier Inverse est donc une collection d'articles dont chacun d'entre eux représente la liste des citations indexées par un descripteur.

A chaque descripteur ou terme d'un lexique (relevant d'un champ pouvant être interrogé) est associé un article dans le fichier Inverse des Descripteurs.

Ce fichier regroupe tous les articles quelque soit l'appartenance des descripteurs aux différents lexiques. Un descripteur peut être introduit dans le lexique sans indexer des documents, il n'y a pas alors d'article dans le fichier Inverse des Descripteurs.

e) Fichier Inverse des Synonymes

Il contient, par classe de synonymes, la liste des documents indexés.

Lorsque deux descripteurs ou plusieurs sont déclarés synonymes, il y a formation d'une classe de synonymie. Si l'un ou plusieurs des descripteurs indexe des documents, il est alors créé dans le fichier Inverse des Synonymes un article qui regroupe la liste fusionnée de l'ensemble des documents indexés par ces descripteurs.

Chaque fichier Inverse est alloué en Unités d'Allocations.

f) Fichier des Liens

Il contient des Liens internes existant entre les fichiers Lexiques, Relations, Inverses des Descripteurs et des Synonymes.

Il assure donc la liaison entre Lexique, fichier Inverse et fichiers Relations.

Comme le lexique, il est subdivisé en catalogues de Liens, un par champ à fichier inverse à chaque descripteur d'un Lexique est associé un numéro de code qui n'est autre que le numéro d'ordre de l'article correspondant dans le catalogue des Liens. Cet article contient quatre pointeurs, à savoir un pointeur sur chacun des fichiers suivants : Lexique (ou Lexique Débordement), Relations, fichier Inverse des Synonymes, fichier Inverse des Descripteurs.

Le pointeur Lexique est toujours présent et permet le retour au lexique à partir du code. Par contre, les autres pointeurs sont nuls si le descripteur n'a pas d'article dans le fichier correspondant.

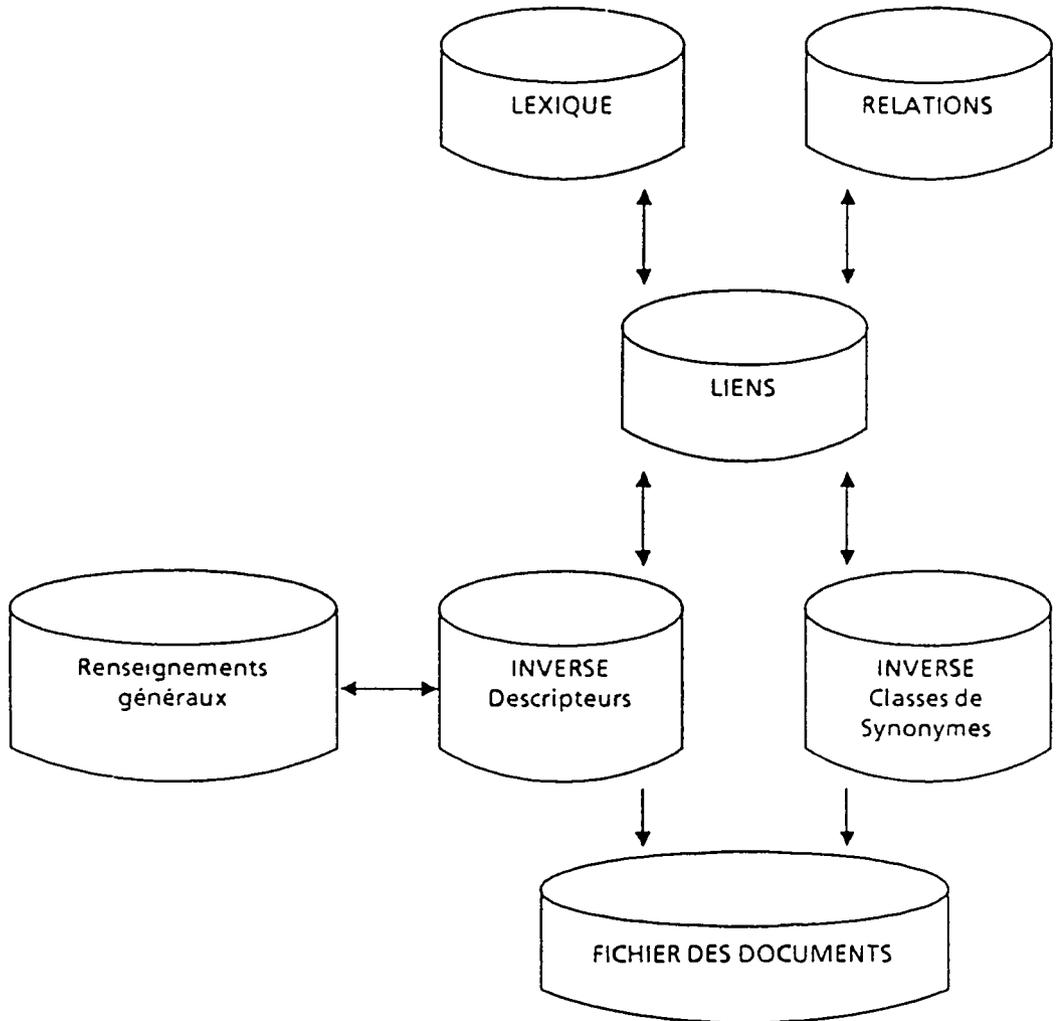
La dimension d'un catalogue de Liens est un nombre entier d'U.A..

* Format d'un fichier Liens

La taille de l'article est de 16 octets :

pointeur lexique	pointeur relations	pointeur F.I. descripteur	pointeur F.I. synonyme
0	4	8	12 16

III-2-3. Organisation générale



Il faut noter que les parties principales des fichiers permanents doivent avoir un nombre entier d'unités d'allocations (1UA = 120 K octets = 120 blocs).

III-3. La création des fichiers permanents d'une base ou CR

Les fichiers permanents d'une base MISTRAL sont créés par cette fonction.

Pour chaque fichier, ADMIST propose une taille en blocs qui peut être modifiée. Puis, il crée le fichier.

Cette fonction ne peut être exécutée qu'une seule fois.

Cette fonction initialise dans le fichier des Renseignements Généraux des informations concernant les fichiers de la base et effectue la création de ces fichiers.

Les fichiers MISTRAL présentent trois états différents :

- état 'fichier inexistant';
- état 'fichier prêt';
- état 'fichier actif'.

Avant la création effective des fichiers de la base, les fichiers ont l'état 'inexistant'. La création les place dans l'état 'prêt'.

La création ne peut se faire que si un mot de passe et une édition cataloguée, au moins, ont été rentrés et que la configuration a été lancée.

Les différentes étapes de la création sont:

- sauvegarde du fichier des renseignements généraux

L'utilisateur peut s'il le veut, sauvegarder son fichier des Renseignements Généraux dans l'état 'avant création'

- utilisation de la saisie

Si l'administrateur n'utilise pas la saisie, il n'y aura pas de création du fichier DOCC (fichier Collection des Documents saisis).

- données sur les fichiers

Les données sont demandées pour les fichiers suivants :

- fichier Sauvegarde des Renseignements Généraux (SARG) s'il y a demande de sauvegarde du fichier des Renseignements Généraux (GINF).
- fichier des Documents principal (DOCM).
- fichier des Documents débordement (DOCO).
- fichier des Lexiques principal (LEXM).
- fichier des Lexiques débordement (LEXO).
- fichier Inverse des Descripteurs (FIDE).
- fichier Inverse des Synonymes (FISY).
- fichier Bibliothécaire (BIBL)(contient les structures, rapports, profils).
- fichier Collection des Documents saisis (DOCC).

Pour le premier fichier créé, SARG s'il y a sauvegarde du fichier des Renseignements Généraux ou DOCM sinon, ADMIST demande le ou les support(s) disque du fichier. Le support entré pour le premier fichier créé est proposé par défaut pour les fichiers suivants. Puis la taille, en blocs, proposée pour le fichier est visualisée.

Cette taille est :

- soit la taille minimale autorisée (500 blocs pour DOCC, 300 pour BIBL);
- soit la taille calculée par le configurateur ou entrée par l'administrateur en mise à jour avant création.

Les tailles calculées par le configurateur pour DOCM, LEXM sont des tailles minimales qui ne peuvent être qu'augmentées. Les tailles des autres fichiers peuvent être modifiées au gré de l'administrateur.

Les tailles de DOCM, LEXM, FIDE, FISY sont multiples de 120 blocs.

Les tailles de DOCO et LEXO sont multiples de 6 blocs.

Pour SARG, la taille est automatiquement de 700 blocs et n'est pas demandée.

La taille maximale du fichier BIBL est de 18432 caractères.

le listing de la création des fichiers permanents pour notre base est donné ci dessous :

fonction ?

cr

sauvegarde du fichier des r.g.(O/N)?

?o

utilisation de la saisie(O/N)?

?o

support du fichier sauvegarde des R.G.:media,devclass

?mist1,ms/d500

support du fichier documents principal:media,devclass

MIST1,MS/D500

?

taille du fichier(en blocs):

6840

modification(O/N)?

+++

?n

support du fichier documents debordemt:media,devclass

MIST1,MS/D500

?

taille du fichier(en blocs):

6

modification(O/N)?

?800

support du fichier lexiques principal :media,devclass
 MIST1,MS/D500
 ?
 taille du fichier(en blocs):
 1800
 modification(O/N)?
 ?n

support du fichier lexiques debordemnt:media,devclass
 MIST1,MS/D500
 ?
 taille du fichier(en blocs):
 354
 +++

taille du fichier(en ^{niveau descripteurs} blocs):
 6840
 modification(O/N)?
 ?n

support du fichier inverse synonymes :media,devclass
 MIST1,MS/D500
 ?
 taille du fichier(en blocs):
 720
 modification(O/N)?
 ?n

support du fichier bibliothecaire :media,devclass
 MIST1,MS/D500
 ?
 taille du fichier(en blocs):
 300
 modification(O/N)?
 ?n

support du fichier descripteurs saisie:media,devclass
 MIST1,MS/D500
 +++

?
 taille du fichier(en blocs):
 500
 modification(O/N)?
 ?700

fonction ?
?vi le

lexiques

+++

```

*****
: indicatif : code : monosens : tail. lx(bl) : tail. ls(bl) :
*****
: REF      : 1 : OUI   : 360 : 240 :
: TIT.     : 2 : OUI   : 600 : 360 :
: AUT.     : 3 : OUI   : 480 : 240 :
: DAT.     : 5 : OUI   : 120 : 120 :
: MCL.     : 6 : OUI   : 120 : 120 :
: CLA.     : 7 : OUI   : 120 : 120 :
*****

```

III-4. Statistiques sur les informations contenues dans les Fichiers Permanents MISTRAL (ST)

Une fois la base créée, l'administrateur peut à tout moment :

- consulter l'état d'un fichier de la base,
- visualiser des statistiques sur son contenu.

Les fichiers pouvant être consultés sont :

- le fichier des Lexiques,
- le fichier des Liens,
- le fichier des Relations,
- le fichier Inverse des Descripteurs,
- le fichier Inverse des Synonymes,
- le fichier des Documents.

Pour chaque fichier, l'administrateur visualise :

- son état :
 - alloué (mais vide),
 - actif (contient des informations),
 - saturé.
- sa taille
- son taux d'utilisation

et pour certains fichiers :

- la taille du fichier débordement, ainsi que certaines informations spécifiques à chaque fichier.

Outre les informations globales, l'administrateur a la possibilité de visualiser des informations détaillées telles que :

- informations lexique par lexique pour la fichier des lexiques,
- informations catalogue par catalogue pour le fichier des liens,

- distributions des fréquences d'indexation pour les fichiers inverse soit sous forme de tableaux, soit sous forme de graphiques.

Quand j'ai créé la base, la plupart des informations ne m'ont pas été données car la base était vide. Les informations que j'ai obtenu sont données ci dessous :

fonction ?

++ (SI)

! To = association de tout d

statistiques sur le fichier lexiques

etat global:

fichier alloue mais vide

taille du fichier principal (en U.A.).....

15

fin des editions statistiques sur ce fichier

statistiques sur le fichier des liens (O/N) ?

?

statistiques sur le fichier des liens

etat global:

fichier alloue mais vide

taille du fichier (en U.A.).....

10

fin des editions statistiques sur ce fichier

statistiques sur le fichier des relations (O/N) ?

?

statistiques sur le fichier relations

fichier alloue mais vide

taille du fichier (en U.A.).....

6

fin des editions statistiques sur ce fichier

statistiques sur le fichier inverse des descripteurs (O/N) ?

?

statistiques sur le fichier inverse des descripteurs

etat global:

fichier alloue mais vide

taille du fichier (en U.A.).....

57

statistiques sur le fichier inverse des synonymes (O/N) ?
?

statistiques sur le fichier inverse des synonymes

etat global:

fichier alloue mais vide

taille du fichier (en U.A.).....

6

fin des editions statistiques sur ce fichier

statistiques sur le fichier des documents (O/N) ?

?

statistiques sur le fichier documents

etat global:

fichier alloue mais vide

taille du fichier principal (en U.A.).....

57

pourcentage de place reservee pour les domaines deja existants.

100

nombre de domaines.....

1

taille du fichier debordement (en blocs).....

804

taux d'utilisation du fichier debordement (%).....

0

nombre total de documents.....

0

fin des editions statistiques sur ce fichier

III-5. Création des premiers formats de saisie

Le programme interactif ADMIST (programme d'administration des bases) permet à l'administrateur de déclarer les différents formats de saisie sur chaque base. En effet, MISTRAL offre une saisie multiformats (50 formats maximum pour une base).

Les formats de saisie servent, bien évidemment, à la création des documents mais ils peuvent être également utilisés pour la modification de documents (comme la mise à jour etc...). Les documents de la base que j'ai créé sont normalement saisis sur DBASEIII mais il est aussi possible de les saisir sur MISTRAL grâce à la grille que l'on va définir lors de la création des formats de saisie.

III-5-1. Principe de fonctionnement logique

Dès le lancement du programme ADMIST, un dialogue s'échange avec l'administrateur des bases (se reporter au paragraphe III appel du programme ADMIST).

Pour la fonction, on entre MJ (Mise à Jour) et l'étape SA (Saisie). A ce niveau, le système propose alors la liste de tous les champs de la base. C'est dans ce lot que l'administrateur puise pour constituer ses différents formats.

D'autre part, les écrans de saisie peuvent être constitués de deux façons différentes. Il est en effet possible de grouper ou non plusieurs champs sur une même page de l'écran de saisie. Ce qui revient à définir deux types de champs pour la saisie :

- les champs "courts" qui contiennent peu d'informations, dont la longueur maximum peut être définie et qui peuvent être groupés sur une même page d'écran.
- les champs "longs", généralement de type texte, dont le contenu est volumineux et non limité et dont le traitement nécessite au moins une page d'écran.

Pour définir un champ de saisie, il faut :

- le nom du champ en abrégé,
- la longueur (en nombre de caractères),
- le type (facultatif ou obligatoire).

Si le format de saisie semble correct à l'administrateur, il le valide. ADMIST (étape SA) propose alors une liste de secret (documents autorisant la saisie de tous les documents (0/119)).

III-5-2. Création des formats de saisie pour notre base.

Tous les formats définis, de même que les formats implicites par domaines sont stockés sur le fichier des Renseignements Généraux de la base considérée.

Les documentalistes ont défini le nombre de lignes que doit avoir chaque champ. Il faudra donc transcrire ces lignes en nombre de caractères sachant qu'une ligne fait 75 caractères et qu'un écran fait 14 lignes.

On détermine donc les champs que l'on visualisera sur chaque écran. J'ai du faire 3 écrans pour pouvoir faire apparaître tous les champs. En effet, le champ RES. (Résumé) doit déjà tenir un écran; c'est un champ long dont la longueur n'est pas défini. Pour les autres champs qui sont donc des champs courts, j'ai essayé de les grouper de manière à avoir des écrans homogènes et logiques.

Le résultat du format de saisie créé est montré ci dessous :

```
S: admist
B U L L  MISTRAL  V5.2-1A
WORK LANGUAGE?
FR  EN
?fr
niveau de traitement?
?ad
numero de la base (entre 1 et 25)?
?13
ouverture des R.G. en modification(MO) ou en visualisation(VI)?
?mc
mot de protection :
```

AXXXXXXXXXXXXX

fonction ?

?m)

visualiser nom de prochaine etape?

?sa)

FORMAT3

TYPE : 1

voici la liste des champs a grouper :

AUT.	BUL.	CHA.	CLA.	DAT.	DIS.
INT.	MCL.	OUV.	RES.	SOU.	TIT.

entrer les champs du groupe : 1

pour chaque champ : nom, longueur, type (fac=@, obl=1)

?tit., 225, @

?aut., 150, @

?sou., 150, @

?mcl., 225, @

?

voici la liste des champs a grouper :

BUL.	CHA.	CLA.	DAT.	DIS.	INT.
OUV.	RES.				

entrer les champs du groupe : 2

pour chaque champ : nom, longueur, type (fac=@, obl=1)

?dat., 8, @

?cla., 4, @

?cha., 75, @

?dis., 1, @

?int., 1, @

?bul., 3, @

?ouv., 1, @

?

voici tous les champs dans le format :

FORMAT3

! champ	! groupe	! longueur max	! type(s)	!
! TIT.	! 0001	! 0225	! indx, msep	!
! AUT.	! 0001	! 0150	! indx	!
! SOU.	! 0001	! 0150	!	!
! MCL.	! 0001	! 0225	! indx	!
! DAT.	! 0002	! 0008	! date, indx	!
! CLA.	! 0002	! 0004	! indx	!
! CHA.	! 0002	! 0075	!	!
! DIS.	! 0002	! 0001	!	!
! INT.	! 0002	! 0001	!	!
! BUL.	! 0002	! 0003	!	!
! OUV.	! 0002	! 0001	!	!
! RES.	! 0003	! ****	!	!

validez-vous ce format (O/N) ?

?o

secret(s) : @:119

Secret des Doc. permis (EX:s1,s2,s3:s4,-s5,-s6:s7 s<120)

?0

Si ce format ne convient pas, on a la possibilité de créer un autre format de saisie.

III-6. Création d'un fichier de mots vides pour les titres

III-6-1. Définition du fichier de mots vides

Pour faire un lexique des mots vides du titre ainsi qu'une liste de mots vides, on utilise l'indexation automatique.

L'indexation automatique consiste à demander à la machine de caractériser elle-même un texte (ici le titre) par des descripteurs.

La méthode d'indexation automatique employée par MISTRAL est la suivante :

- le documentaliste fournit en donnée le texte à analyser.
- MISTRAL lit ce texte, en extrait les mots et supprime de ceux-ci, ceux qui sont sans signification ou qui n'apportent rien à la compréhension du texte, c'est à dire les mots vides.

Ces mots vides sont entre autre :

- les articles,
- les prépositions,
- les verbes trop généraux.

Les mots obtenus finalement sont considérés comme des descripteurs et sont associés à un fichier Inverse.

Par exemple, quand je donne un titre à indexer à MISTRAL, MISTRAL va d'abord lire le texte et en extraire tous les mots.

Un 'mot' correspond à tout groupe de lettres séparé par des blocs ou des signes de ponctuations diverses.

Ces mots sont ensuite comparés à la liste des mots vides que l'on a faite. Si un mot appartient à cette liste, il est éliminé. Les termes restant sont alors considérés comme des descripteurs.

Limite

- il est possible de définir jusqu'à 5 fichiers mots vides pour traiter des champs dont on demande l'indexation automatique.
- le nombre maximum de caractères par fichier est de 6104.
- les mots vides doivent avoir une longueur inférieure ou égale à 20 caractères. Au delà, les mots sont tronqués.
- on ne peut modifier un fichier de mots vides existant sans risque d'incohérence entre le fichier des Documents et le lexique. Ce type de traitement n'est donc pas permis.

III-6-2. Cr ation du fichier mot vide

Les titres du dossier de presse sont soit en franais, soit en anglais. Donc, notre fichier mot vide comprendra des mots vides en franais et en anglais.

1) J'utilise comme r f rence une liste de mots vides d'une base d j  cr e e. Cette liste est alors copi e sur notre base. Le programme de ce transfert est donn  ci dessous :

```
(1)      S : LMN SL .V5DON
(2)      C : FSE
(3)      F : MD B12-MOTVID
(4)      : SVQ * B13-MOTVID
(5)      F : /
(6)      C : PR B13-MOTVID
```

explication des commandes :

- (1) les commandes ou param tres pour chaque programme sont contenus dans un membre de la biblioth que .V5DON en format ' diteur de texte' qui est assign    l' tiquette ENTREE.
- (2)  diteur pleine page.
- (3) cette commande nous donne la liste des mots vides de la base 12 qui servira de mod le et qui sera donc transf r e sur notre base.
- (4) sauvegarde de la liste des mots vides de la base 12 sur notre base 13.
- (5) on veut remonter   l' tape pr c dente qui est C.
- (6) impression de cette liste.

2) J'ai ensuite analys  les titres de 10 bulletins de signalisation des articles de presse que la documentaliste m'a envoy s. Cette analyse m'a permis de rajouter des mots vides   la liste. On a aussi la possibilit  d' liminer certains mots vides mais cela n'a pas un grand int r t.

Le programme pour ajouter ou effacer des mots vides est :

```
(1)      F : AD B13-MOTVID
(2)      F : A 130
(3)      F : D 140
```

explication des commandes

- (1) commande pour pouvoir modifier la liste des mots vides de notre base.

(2) commande qui permet d'ajouter un ou plusieurs mots vides.
Ici l'ajout se fera à la ligne 130.

(3) commande qui permet de détruire le mot vide de la ligne
140.

3) La plupart des titres sont composés de verbes. Il y a donc fallu que je définisse les verbes qui feraient partie de la liste des mots vides. Il était hors de question de les conjuguer tous.

Avec la documentaliste, on a décidé d'insérer dans cette liste les verbes trop généraux tel que avoir, être, have, be, ainsi que ceux qui apparaissent dans les titres fréquemment. Les autres verbes apparaîtront alors dans le lexique des mots du titre.

La liste des mots vides obtenue est donnée en annexe 2.

III-6-3. Modification des séparateurs (SP)

Les séparateurs ont tous un seul caractère alphanumérique. Pour chaque groupe, une liste de valeurs est proposée par défaut, en représentation interne et alphanumérique.

On veut modifier la liste des séparateurs du champ titre.

En effet, certains titres contiennent des mots composés (comme sous-marin, war's) qui quand ils seront indexés dans le lexique titre sortiront comme 2 mots clés, ce qui nécessiterait lors d'une interrogation de combiner les 2 mots clés entre eux. On veut donc faire apparaître dans le lexique du titre, le mot composé en entier c'est à dire qu'il formera un seul et unique mot clé.

Le champ titre est un champ multiséparateur donc la table des séparateurs contient le caractère blanc et tous les caractères de ponctuation.

J'ai donc éliminé les séparateurs "-" et "'". Ces séparateurs iront dans la liste des mots vides.

ADMIST visualise les valeurs des séparateurs en code interne et alphanumérique.

La modification et le résultat sont montrés ci dessous :

reprandre,visualiser ou prochaine etape (RE,VI nom1,nom2)?

SP

séparateurs de champs multiseparateurs

H40 H4A H4B H4C H4D H4E H50 H5A H5B H5C H5D H5E H61 H6B H6D H6E

° . < (+ & § \$ *) ; / , - >

H6F H7A H7B H7D H7E H7F HE0

? : £ ' = " ç

separateurs a retirer(liste ou DE seulement pour rel.auto.)

?_'

caracteres speciaux a retirer(en format interne-1 par ligne)

?

liste des separateurs a ajouter

?

entree des caracteres speciaux(en format interne-1 par ligne)

?

```

separateurs de champs a relations automatiques
H4@ H4B H4D H5C H5D H5E H6@ H7A H7D
. ( * ) ; - : '
separateurs a retirer(liste ou DE seulement pour rel.auto.)
?
caracteres speciaux a retirer(en format interne-1 par ligne)
?
liste des separateurs a ajouter
?
entree des caracteres speciaux(en format interne-1 par ligne)
?

reprendre,visualiser ou prochaine etape (RE,VI nom1,nom2)?
?vi sp

```

+++

```

separateurs de champs multiseparateurs
*****
H4@ H4A H4B H4C H4D H4E H5@ H5A H5B H5C H5D H5E H5I H6B H6E H6F H7A
. ( + & @ # * : ; / , : ? :
H7B H7E H7F HE@
E = " ¢

separateurs de champs a relations automatiques
*****
H4@ H4B H4D H5C H5D H5E H6@ H7A H7D
. ( * ) ; - : '

```

Il faut ensuite ajouter dans la liste des mots vides les séparateurs "-" et "'" qui ont été retirés du champ multiséparateur du lexique titre.

L'incorporation de ces 2 séparateurs se fait avec la commande A (Ajout) (cf explication plus haut). Il faut tenir compte lors de l'ajout de ces caractères de l'ordre imposé. En effet, la liste des mots vides doit d'abord commencer par les minuscules, puis les majuscules et enfin les séparateurs. Ils faut donc les mettre à la fin de la liste.

Un extrait de la liste des mots vides montre l'incorporation des séparateurs..

S:??

TOP

```

>>>>>10 IDEN KR

      20 FARM INIT

      30 FLD TIT.
      40 A

      40.1 AINSI
      50 AN

      60 AND

      60.1 ANNUAL
      60.2 APRES
      60.3 ARE
      70 AS
      80 AT
      90 AU
      100 AUX
          1 $

      810 Y
      820 Z
          1 -
          2 '

```

IV. Sauvegarde de la base

Avant de commencer à alimenter la base, on sauvegarde la base "vierge" sur bande; ceci afin de pouvoir faire un retour arrière si au premier chargement de la base on s'aperçoit d'anomalies.

V. MISE EN MARCHE ET INTERROGATION

Les documents saisis sur DBASEIII sont alors transférés du micro sur le central. Donc, avant toute interrogation sur MISTRAL, il faut passer la chaîne de Mise à Jour.

Pour faire une interrogation, il faut utiliser des procédures de recherche qui se distinguent des autres messages par le fait qu'elles débutent toujours par les deux caractères M : (suivi du nom de la procédure).

Il existe plusieurs types de recherche :

- recherche sur références,
- recherche par intervalles de valeur,
- recherche sur fichier inverse,
- recherche par auteur.

V-I. Recherche sur références

Pour pouvoir effectuer cette interrogation, il convient de connaître la référence de l'article.

Dans notre base, nous avons défini une référence comme ayant une longueur maximum de 8 chiffres. Si cette longueur est inférieure, MISTRAL la complète alors avec des blancs.

Syntaxe

M : QU[ESTION] RF { référence de document }
 { liste de références }

paramètres

- RF : sigle de commande spécifiant une recherche par référence de documents.
- Référence de document : chaîne de caractères.
- Liste de références : référence 1, ..., référence n.

Toute référence inconnue est signalée.

Quand le nombre de références répondant au critère de sélection dépasse 100, seuls les 100 premières références sont retenues pour constituer la liste résultat.

Exemple

Je veux rechercher la référence 700571 dans la base créée.
 La procédure de recherche est donnée ci-dessous :

```
Procédure ou étape de recherche 1
?m:qu rf 700571
*** 1*** resultat:      1
```

Pour visualiser la référence demandée, on utilise la procédure suivante :

```
Procédure ou étape de recherche 2
?m:vi ecl
```

```
REF      : 900571
TIT.     : ELECTRICITY BY SERENDIPITY
AUT.     : POOL R
SOU.     : SCIENCE
DAT.     : 90/03/02
MCL.     : FILE, HYDROGENE, OXYGENE, GAZ
```

EC1 correspond à l'édition cataloguée 1 définie lors de la fabrication de la base. Cette édition ne permet que la visualisation de certains champs.
En utilisant l'édition cataloguée 2 EC2, on visualise tous les champs.

Procédure ou étape de recherche 2
?m:vi ec2

REF : 900571
TIT. : ELECTRICITY BY SERENDIFITY
AUT. : POOL R
SOU. : SCIENCE
DAT. : 90/03/02
MCL. : PILE, HYDROGENE, OXYGENE, GAZ
CLA. : XXII
CHA. : INFORMATIONS DIVERSES
DIS. : Ø
RES. : LA NOUVELLE PILE, DECOUVERTE PAR HASARD PAR C. DYER, PRODUIT DE L'ENERGIE ELECTRIQUE A PARTIR D'UN MELANGE HYDROGENE/OXYGENE, EVITANT LES COMPLICATIONS APORTEES PAR LES DEUX SYSTEMES SEPARES D'ARRIVEE DE GAZ. LE FONCTIONNEMENT EST DU A "UN NOUVEAU TYPE DE CATALYSE" QUI EST "PLUS PRES DE LA BIOCHIMIE QUE DE LA CHIMIE CLASSIQUE".

Procédure ou étape de recherche 2
?

V-2. Recherche par intervalle de valeur

La recherche sur intervalle de valeur permet de chercher les documents contenant, dans un champ donné (numérique ou date) une valeur comprise dans un intervalle donné.

Syntaxe

M : QU[ESTIQN] QI/champ[caractéristique] } > borne } { > borne {
 } < } { < }

paramètres :

- M : QU : cette commande n'est pas obligatoire.
- QI : sigle issu du fichier dialogue spécifiant une recherche sur intervalle.
- /champ : nom du champ alimentant un lexique.
- { } : symbole mathématique caractérisant la valeur de la
 } < } > borne inférieure et supérieure de l'intervalle.
- borne : valeur alphanumérique.

Exemple

Soit la base dossier de presse dans laquelle les documents possèdent un champ "DAT.", contenant l'année de parution du document. La recherche par intervalle de valeur sur ce champ permet de choisir des documents parus entre telle année et telle autre année.

Je vais donc chercher les documents qui sont parus en 1989.

La procédure de recherche est donnée ci-dessous :

```

Procédure ou étape de recherche 3
?qi/DAT. >1988<1990
*** 3*** resultat:      1
Procédure ou étape de recherche 4
?m:vi

```

```

REF      : 900009
TIT.     : SELON L'ANNUAIRE "FLOTTES DE COMBAT 1990", LA MARINE SOVIETIQUE
          N'A PAS VERITABLEMENT DESARME
SOU.     : MONDE
DAT.     : 89/12/26
MCL.     : URSS, SNA, SNLE

```

V-3. Recherche sur fichier inverse

Cette recherche effectuée la sélection d'un ensemble de documents à travers les lexiques et fichiers inverses.

Syntaxe

On entreprend une recherche en spécifiant des termes isolément ou en les combinant avec les opérateurs booléens (ET, OU, SAUF). La réponse du système indique le nombre de documents contenant les descripteurs correspondant. Cette méthode peut être répétée autant de fois que nécessaire, par combinaison des résultats antérieurs entre eux ou avec de nouvelles intégrations.

Exemple

J'ai effectué deux interrogations en utilisant les opérateurs booléens et les étapes de recherche antérieures.

```

Procédure ou étape de recherche 4
?
?laser ou chimie
*** 6*** resultat:      3
Procédure ou étape de recherche 7
?6 et optique
*** 7*** resultat:      1
Procédure ou étape de recherche 8
? n:vi Ec2

```

REF : 900042
 TIT. : ULTRAFAST DYNAMICS AT SEMICONDUCTOR AND METAL SURFACES
 AUT. : BOKOR J
 SOU. : SCIENCE
 DAT. : 89/12/09
 MCL. : LASER, OPTIQUE, CHIMIE
 CLA. : XXII
 CHA. : INFORMATIONS DIVERSES
 DIS. : 0
 RES. : LES MESURES ULTRARAPIDES UTILISENT LASERS ET OPTIQUES NON
 LINEAIRES. LES CARACTERISTIQUES DE RELAXATION DES EXCITATIONS OPTIQUES
 EN SURFACE ET CELLES DES CORPS ABSORBES DONNENT UN NOUVEL APERCU SUR LA
 CHIMIE DE SURFACE, LES TRANSITIONS DE PHASE ET LA RECOMBINAISON DES
 PORTEURS DE CHARGE DANS LES SEMI-CONDUCTEURS.

2ème exemple :

Procedure ou etape de recherche 8
 ?

?hydrogene

*** 8*** resultat: 3

Procedure ou etape de recherche 9

?(8 ou gaz) sauf oxygene

*** 9*** resultat: 2

Procedure ou etape de recherche 10

?m:vi

REF : 900101
 TIT. : DES MIROIRS A ATOMES
 SOU. : RECHERCHE
 DAT. : 90/01/00
 MCL. : HYDROGENE, MIROIR, CONCAVE

REF : 900337
 TIT. : L'HYDROGENE SOUS PRESSION : DU GAZ AU METAL
 AUT. : MONAT B
 SOU. : RECHERCHE
 DAT. : 90/03/00
 MCL. : HYDROGENE, GAZ, SPECTRE, EMISSION

V-4. Recherche par auteur

Cela permet de retrouver des documents à partir du nom de l'auteur. Il faut toujours indiquer le champ suivi du nom de l'auteur.

Je veux par exemple rechercher le document qui a été écrit par CADIOU JM.

Procédure ou étape de recherche 10

? :aut. cadiou jm

10 resultat: 1

Procédure ou étape de recherche 11

?m:vi ec2

REF : 900606

TIT. : L'EUROPE A REDONNE CONFIANCE A SES CHERCHEURS

AUT. : CADIOU JM

SOU. : 01 INFORMATIQUE

DAT. : 90/04/06

MCL. : ESPRIT, INFORMATION, PROGRAMME

CLA. : VII

CHA. : EUROPE-OTAN

DIS. : 0

RES. : LE PROGRAMME ESPRIT (TECHNOLOGIES DE L'INFORMATION) A NOTAMMENT PERMIS LA CONSTRUCTION D'UNE VERITABLE EUROPE DES CHERCHEURS ET REDORE LE BLASON TECHNOLOGIQUE DU VIEUX CONTINENT.

Procédure ou étape de recherche 11

?

AUTRES POSSIBILITES
DE MISTRAL

TROISIEME PARTIE

I. LA SAISIE

Cette étape ne peut être appelée qu'après la création des fichiers de la base en mise à jour après création. La fonction de saisie est complémentaire de la fonction d'interrogation du système interactif MISTRAL. Il est clair que l'utilisateur ne peut activer la procédure de saisie que si l'habilitation le lui permet.

I-1. Les fonctionnalités du système

Le module de saisie assure essentiellement les quatre fonctions suivantes :

- création de document
- modification de document
- retrait d'un document
- visualisation d'un document.

Selon la fonction choisie, le système affiche sur le terminal successivement différents types de grille qui s'enchainent suivant un ordre déterminé.

I-1-1. Création de documents

Grâce à l'utilisation de formats de saisie défini par l'administrateur, l'utilisateur est guidé pas à pas dans l'introduction des champs. Il valide chaque champ ou groupe de champs selon la technique de groupage utilisée.

Les champs alimentant un lexique sont analysés par MISTRAL et découpés en descripteurs grâce aux espaces, signes de ponctuation et signes spéciaux.

A la fin de la saisie, l'utilisateur le valide ou revient sur certaines zones pour correction.

A la validation du document, celui-ci est immédiatement introduit dans la base.

I-1-2. Modification - mise à jour de documents

L'utilisateur peut travailler sur les documents résultants de la dernière question ou accéder directement aux documents par leur référence, pourvu qu'il en ait l'habilitation.

L'utilisateur peut alors :

- modifier
- insérer un champ
- supprimer

En fin de mise à jour, il valide le traitement champ par champ. Quand tout est terminé pour un document, il valide le traitement pour ce document. Celui-ci remplace immédiatement l'ancien dans la base.

I-1-3. Suppression de documents

Le traitement peut s'effectuer sur :

- des documents résultants d'une question;
- des documents accessibles par leur référence.

I-1-4. Visualisation des documents

Cette visualisation peut être demandée pour un document existant :

- avant tout traitement de modification pour un document existant;
- à la fin du traitement du document en création ou en modification, avant d'en demander la validation.

I-2. Activation du module de saisie/mise à jour

Il faut lancer le programme MISSAI.

Avant d'utiliser la procédure M:SA (SAisie), l'utilisateur doit :

- obligatoirement être connecté, implicitement ou explicitement (procédure M:BA) à la base qui doit être mise à jour.
- spécifier le nom du domaine concerné M:DO;
- fournir sa clé d'habilitation. Cette clé d'habilitation est fournie en réponse à la question 'nom de votre clé de secret?' après avoir activée la procédure M:DE.
- enfin, on peut mettre la procédure M:SA.

Cette procédure provoque le passage d'un dialogue en mode télétype (ligne par ligne), qui est le mode standard de MISTRAL pour la consultation des bases, à un dialogue en mode page. Ce dernier entraîne l'affichage d'écrans qui se substituent l'un à l'autre chaque fois que l'on appuie sur la touche 'TRANSMIT'.

IDENTIFICATION--

AXXXXXXXXXXXXXX

Base Connectee: BASE002

Procédure ou etape de recherche 1

?m:ba

BASE EN LIGNE

1. BASE001 BASE 01
2. BASE002 BASE 02
3. BASE003 BASE 03
4. BASE004 BASE 04
5. BASE005 BASE 05
6. BASE006 BASE 06
7. BASE007 BASE 07
8. BASE008 BASE 08
9. BASE009 BASE 09
10. BASE010 BASE 10
11. BASE011 BASE 11

+++

12. BASEØ12 BASE 12
 13. BASEØ13 BASE 13
 Selectionner ou non ?
 ?13

Base Connectee: BASEØ13
 Procedure ou etape de recherche 1
 ?
 ?m:do presse
 PRESSE Mot de passe du domaine ?

AYXXXXXXXXXX

Procedure ou etape de recherche 1
 ?m:de
 Nom de votre cle de secret ?

AXXXXXXXXXXX

Procedure ou etape de recherche 1
 ?m:sa
 ?

I-3. Les différents types de grilles nécessaires pour la création d'un document

Le système gère 5 types de grille :

GRILLE 1 - Choix des options générales de traitement.
 Cette grille est la première qui s'affiche après avoir spécifié la procédure M:SA et c'est toujours celle qui est réaffichée quand on a achevé le traitement d'un document.

GRILLE 2 - Traitement d'un groupe de champs 'courts'.

GRILLE 3 - Traitement d'un champ 'long'.

GRILLE 4 - Fonction éditeur de texte.

Cette grille apparaît à l'écran chaque fois que dans la grille 3 l'on a spécifié l'ordre ST (STRING).

GRILLE 5 - Choix des options en fin de traitement.

Il s'agit de la dernière grille qui s'affiche lors de la création ou de la modification d'un document.

Elle offre les 4 options suivantes :

- * validation du traitement
- * annulation du traitement

- * visualisation du document qui vient d'être traité
- * modification du document qui vient d'être créé ou qui vient de subir une modification.

I-4. Description détaillée des grilles

Toutes les grilles utilisées par le système de Saisie/Mise à jour sont constituées de 21 lignes de 80 caractères. La 21ème ligne est réservée aux messages émis par le système.

Ces messages sont de 2 types :

- message d'erreur;
- message de contrôle précisant le traitement qui a été effectué à partir de la grille affichée sur l'écran.

I-4-1. GRILLE 1 : Choix des options générales de traitement

BASE BASE013

MISTRAL

___: TYPE DE TRAITEMENT

CR : CREATION
 RE : RETRAIT
 MO : MODIFICATION
 VI : VISUALISATION
 FI : FIN DE LA SAISIE
 FD : FIN DU DOCUMENT
 RP : REPRISE DU DOCUMENT

DOCUMENT :

SECRET (0-119) :

CHAMP (SOS=??) :

DOMAINE (SOS=??) :

PRESSE

FORMAT (SOS=??) :

FORMAT3

1) TYPE DE TRAITEMENT

Il est choisi par l'utilisateur.

On doit toujours remplir cette rubrique en mettant un des 7 sigles de la liste affichée (2 caractères).

2) DOCUMENT

On doit donner la référence du document. C'est une zone de 8 à 34 caractères dont la longueur est fixe pour une base donnée. Cette zone doit aussi être obligatoirement remplie.

3) SECRET

C'est le niveau de confidentialité.

Cette zone n'est pas obligatoirement renseignée.

Si celle n'est pas renseignée, cela implique :

- * en création de document que le secret est égal à zéro.
- * en modification que le secret initialement attribué au document est conservé.

4) CHAMP

Cette zone n'est renseignée qu'en cas de modification de document si la modification porte sur un champ particulier.

5) DOMAINE

En création de document, l'utilisateur a 3 possibilités :

- * ou il spécifie le domaine de son choix,
- * ou il laisse cette zone à blanc et c'est alors le domaine implicite qui sera pris en compte,
- * ou il tape la commande d'assistance (??), ce qui provoque l'affichage de la liste des formats implicites par domaines.

5) FORMAT

Si on ne met rien dans cette zone, c'est le format implicite du domaine traité qui est pris en compte. Si on tape la commande d'assistance (??), on obtient la liste des formats disponibles pour la base courante.

Je veux mettre un nouveau document dans la base, je vais donc utiliser le sigle CR pour création et j'inscris la référence du document.

BASE BASE013

MISTRAL

cr : TYPE DE TRAITEMENT

CR : CREATION
 RE : RETRAIT
 MO : MODIFICATION
 VI : VISUALISATION
 FI : FIN DE LA SAISIE
 FD : FIN DU DOCUMENT
 RP : REPRISE DU DOCUMENT

DOCUMENT :

900606

SECRET (0-119) :

CHAMP (SOS=??) :

DOMAINE (SOS=??) :

FRESSE

FORMAT (SOS=??) :

FORMAT3

I-4-2. GRILLE 2 : Traitement d'un groupe de champs 'courts'.

Dans notre base, les champs courts sont dans le groupe 1 et sont présentés comme je l'ai défini dans le programme ADMIST. On verra donc apparaître 2 écrans de champs courts. A l'écran, il suffit de remplir les champs demandés.

En bas de l'écran, 4 options sont offertes à l'utilisateur :

- 1 VALIDER
- 2 ANNULER
- 3 VALIDER ET FIN
- 4 ANNULER ET FIN

Il suffit d'inscrire le numéro choisi.

Les champs courts du document que je crée sont donnés ci-dessous.

CREAT BASE013 ;PRESSE RF= 900606 GROUPE: 0001

TIT. : L'Europe a redonne confiance a ses chercheurs

AUT. : Caciou jm

SOU. : 01 informatique

MCL. : escrit, information, programme

 COMMANDE?: 1 1:VALIDER ; 2:ANNULER ; 3:VALIDER ET FIN ; 4:ANNULER ET FIN

CREAT BASE013 ;PRESSE RF= 900606 GROUPE: 0002

DAT. : 90/04/05

CLA. : VII

CHA. : europe-otax

DIS. : 0

INT. :

BUL. :

OUV. :

 COMMANDE?: 1 1:VALIDER ; 2:ANNULER ; 3:VALIDER ET FIN ; 4:ANNULER ET FIN

I-4-3. GRILLE 3 : Traitement d'un champ 'long'

On observe en haut de l'écran 13 ordres que l'on peut donner à l'éditeur de texte qui gère les champs longs. Ces 13 ordres sont expliqués dans le chapitre "commandes de modification des champs longs".

La zone texte comprend 16 lignes de 78 caractères.

```

CREAT REE.                RF= 900606                PRESSE BASE013
ORDRE: SV                 BA,RE,VI,SU,IN, //          ST
PAGE : _____        EC,VC,SV,RV                 ST      FI
-----

```

-----DERNIERE PAGE :0000

Le programme ESPRIT (technologies de l'information: a notamment permis la construction d'une véritable Europe des chercheurs et redonne le blason technologique du vieux continent.

I-4-4. GRILLE 4 : Fonction éditeur de texte

L'explication de cette grille est faite dans le chapitre "commandes de modification des champs longs".

I-4-5. GRILLE 5 : Choix des options en fin de traitement

C'est le point de passage obligatoire en fin de traitement d'un document.

L'opérateur a le choix entre 4 solutions référencées chacune par un chiffre :

1. VALIDATION DU DOCUMENT
2. ANNULATION DU TRAITEMENT
3. VISUALISATION
4. MODIFICATION

En fonction de l'option choisie, l'utilisateur renseigne la zone indiquée FIN DE TRAITEMENT DU DOCUMENT.

J'ai donc utilisé l'option 1 ce qui valide le document créé.

```

CREATION          BASE :BASE013
DOMAINE           :PRESSE
DOCUMENT          :900606
SECRET           :00

```

```

FIN DE TRAITEMENT DU DOCUMENT: 1
      1 : VALIDATION DU TRAITEMENT
      2 : ANNULATION DU TRAITEMENT
      3 : VISUALISATION
      4 : MODIFICATION

```

```

REF      : 900606
TIT.     : L'EUROPE A REDONNE CONFIANCE A SES CHERCHEURS
AUT.     : CADIOU JM
BOU.     : 01 INFORMATIQUE
MCL.     : ESPRIT, INFORMATION, PROGRAMME
DAT.     : 90/04/06
CLA.     : VII
CHA.     : EUROPE-OTAN
DIS.     : 0
RES.     : LE PROGRAMME ESPRIT ( TECHNOLOGIES DE L'INFORMATION) A
          NOTAMMENT PERMIS LA CONSTRUCTION D'UNE VERITABLE EUROPE DES CHERCHEURS
          ET REDORE LE BLASON TECHNOLOGIQUE DU VIEUX CONTINENT.

```

Puis le système réaffiche alors la grille 1 prête pour un autre traitement. Si on veut arrêter, il suffit de mettre les caractères FI (Fin).

BASE BASE013

MISTRAL

fi : TYPE DE TRAITEMENT

```

CR : CREATION
RE : RETRAIT
MO : MODIFICATION
VI : VISUALISATION
FI : FIN DE LA SAISIE
FD : FIN DU DOCUMENT
RF : REPRISE DU DOCUMENT

```

DOCUMENT :

SECRET (0-119) :

CHAMP (SOS=??) :

DOMAINE (SOS=??) :

PRESSE

FORMAT (SOS=??) :

I-5. Règles à savoir pour la saisie

* Avant toute saisie et interrogation, il convient de connaître les règles de syntaxe des opérateurs et des caractères de fonctions.

* On peut saisir sur MISTRAL en majuscule ou en minuscule. Lors de la visualisation du document final, MISTRAL transcrit tout en majuscule.

Si certains champs ne sont pas remplis, ils n'apparaissent pas lors de la visualisation du document total.

* Pour les champs longs c'est à dire ici pour le résumé, si on saute une ligne entre 2 paragraphes, celle-ci disparaît quand on visualise le document en entier. Par contre, si on va à la ligne pour créer un paragraphe, MISTRAL met la phrase à la suite en laissant des blancs entre les 2 phrases. Il faudra alors utiliser la fonction ST (STRING) pour éliminer cet espace.

* Enfin, il faut aussi soulever le problème des mots creux et des blancs.

Ce problème doit bien être défini au début lors de la création de la base et avant la saisie. En effet, la plupart des bruits qui apparaissent quand on fait une interrogation proviennent de cela.

Donc, si entre le tiret qui sépare 2 mots formant un mot composé, on met un blanc, le tiret sera considéré comme un mot vide.

Exemple : Etats - Unis

On aura donc ici 2 mots clés séparés qui sont Etats et Unis.

Par contre, si on écrit Etats-Unis, ce mot sera considéré comme un seul mot clé. Il faut donc faire très attention lors de la saisie à la présence ou non de blanc.

De plus, il faudra séparer chaque mot clé par une virgule sinon ils seront considérés comme un seul mot clé. Cela est de même pour les auteurs. En effet, si on oublie les virgules séparatrices, il n'y aura pas de lexique possible.

II. COMMANDES DE MODIFICATION DES CHAMPS LONGS

La commande STRING (ST) déclenche l'affichage d'une grille à partir de laquelle on peut, grâce à un éditeur de texte :

- remplacer, insérer ou retirer une chaîne de caractères;
- comprimer des blancs.

En effet, en création ou plus souvent en modification de documents, l'utilisateur peut être amené à vouloir effectuer sur les champs de type 'long' comme les résumés des traitements sur des chaînes de caractères.

La méthode de travail consiste à choisir une fonction, à renseigner les bornes et chaînes pour effectuer le traitement des chaînes de caractères, puis en fin d'initialisation des fonctions STRING à choisir la fonction 5 "FIN" qui renvoie au résumé avec la modification faite.

L'exemple ci-dessous présente une notice créée sur MISTRAL et dont le résumé a besoin d'être modifié.

```

REF.      : 900571
TIT.      : ELECTRICITY BY SERENDIPITY
AUT.      : POOL P
SOU.      : SCIENCE
MOT.      : PILE, HYDROGENE, OXYGENE, GAZ
DAT.      : 90/03/02
CHA.      : XXII
CHA.      : INFORMATIONS DIVERSES
DIS.      : 0
RES.      : LA NOUVELLE PILE, DECOUVERTE PAR C. DYER, PRODUIT DE L'ENERGIE
          : ELECTRIQUE A PARTIR D'UN MELANGE HYDROGENE/OXYGENE, EVITANT AINSI LES
          : COMPLICATIONS APPORTEES PAR LES DEUX SYSTEMES SEPARES D'ARRIVEE D'AIR.
          : LE FONCTIONNEMENT EST DU A "UN NOUVEAU TYPE DE CATALYSE" QUI EST "PLUS
          : PRES DE LA BIOCHIMIE QUE DE LA CHIMIE CLASSIQUE".

```

Pour pouvoir utiliser la fonction STRING ou ST, il faut lancer la commande MO (MODIFICATION) du document (cf écran ci-dessous)

```

BASE      BASE013                                MISTRAL

MO : TYPE DE TRAITEMENT
  CR : CREATION
  RE : RETRAIT
  MO : MODIFICATION
  VI : VISUALISATION
  FI : FIN DE LA SAISIE
  FD : FIN DU DOCUMENT
  RP : REPRISE DU DOCUMENT

DOCUMENT :
900571
SECRET (0-119) :
CHAMP (SOS=??) :
DOMAINE (SOS=??) :
PRESSE
FORMAT (SOS=??) :
FORMAT3

```

Il faut ensuite valider les champs courts. Quand on arrive au résumé (champ long), dans le champ ORDRE, on met la commande ST (cf exemple ci-dessous).

```

MODIF RES.      RF= 900571                                PRESSE BASE013
ORDRE:  st      SA.RE.VI.SU.IN.//                          ??
PAGE :  1      EC,VC,SV,RV                                ST      FI
PAGE  0001-----DERNIERE PAGE 1001
LA NOUVELLE PILE, DECOUVERTE PAR C. DYER, PRODUIT DE L'ENERGIE ELECTRIQUE A PA
RTIR D'UN MELANGE HYDROGENE/OXYGENE, EVITANT AINSI LES COMPLICATIONS APPORTEES
PAR LES DEUX SYSTEMES SEPARES D'ARRIVEE D'AIR.
LE FONCTIONNEMENT EST DU A "UN NOUVEAU TYPE DE CATALYSE" QUI EST "PLUS PRES DE
LA BIOCHIMIE QUE DE LA CHIMIE CLASSIQUE".

```

N.B. Les autres commandes possibles sont :

- SA = SAUVER
Cette commande provoque la sauvegarde du texte saisi pour constituer une page qui vient s'ajouter à la dernière page du champ.

- RE = REMPLACER
Commande utilisée pour la modification d'une page. Son exécution entraîne le remplacement de la page dont le numéro est spécifié dans la zone de communication.

- VI = VISUALISER
Cette commande provoque l'affichage de la page spécifiée.

- SU = SUPPRIMER
Cette commande permet de supprimer une page d'un champ paginé comportant plusieurs pages.

- IN = INSERER
Cet ordre permet d'insérer une page parmi des pages préexistantes.

- // = DECOUPER
Cette commande permet de découper une page en plusieurs pages de façon à libérer la place nécessaire à l'insertion de plusieurs lignes.

- EC = EFFCHP
Cette commande annule le traitement d'un champ long en cours de création ou de modification.

- VC = VALCHP
Cet ordre permet de valider le traitement du champ. Il est nécessaire pour valider le processus de création ou de modification du champ long en cours de traitement à moins que l'on utilise les commandes SV ou RV.

- SV = SAUVER et VALIDER
Cette commande est équivalente aux ordres SA et VC exécutés consécutivement.

- RV = REMPLACER et SAUVEGARDER
Cette commande est équivalente à l'ordre RE immédiatement suivi de l'ordre RV. Elle permet donc de sauver une page que l'on vient de modifier et de valider le champ modifié.

- FI = FIN
Cet ordre annule la création ou la modification du champ en cours de traitement, et permet en outre d'aller à la fin du traitement du document.

On voit alors apparaître un écran nous donnant les différentes fonctions disponibles sur la commande ST.

```

MODIF RES.                RF= 900571                PRESSE BASE013

FONCTION STRING SUR LE CHAMP ENTIER:
    1 : INSERTION          2 : SUPPRESSION          3 : REMPLACEMENT
    4 : COMPRESSION       5 : FIN
    6 : SUPPRESSION ENTRE 2 BORNES, BORNES COMPRISES

CHAINE-1 :
CHAINE-2 :
BORNE-1  :
BORNE-2  :

```

II-1. Insertion de chaînes de caractères

Zones variables à renseigner :

- CHAINE 1 : chaîne de caractères derrière laquelle a lieu l'insertion.
- CHAINE 2 : chaîne à insérer. Cette chaîne peut comporter un ou plusieurs espaces en tête de chaîne : ils seront insérés.
- BORNE 1 : chaîne de caractères délimitant dans le champ le début de la fenêtre de travail. Si cette chaîne est vide, c'est le début du champ.
- BORNE 2 : chaîne de caractères délimitant la fin de la fenêtre. Si cette chaîne est vide, c'est la fin du champ.

Un message en bas de l'écran indique combien de fois le traitement a été effectué sur le champ.

Sur notre exemple, je veux insérer après "découverte", le groupe de mots "par hasard".

```

MODIF RES.                RF= 900571                PRESSE BASE013

FONCTION STRING SUR LE CHAMP ENTIER:   1
    1 : INSERTION          2 : SUPPRESSION          3 : REMPLACEMENT
    4 : COMPRESSION       5 : FIN
    6 : SUPPRESSION ENTRE 2 BORNES. BORNES COMPRISES

CHAINE-1 : DECOUVERTE
CHAINE-2 : PAR HASARD
BORNE-1  :
BORNE-2  :

```

MESSAGE: NB D'OCCURENCES EFFECTUEES PAR CETTE FONCTION : 000001

La commande FIN (5) réaffiche le résumé avec le nouveau texte entré.

```

MODIF RES.                RF= 900571                PRESSE BASE013
DEPRE: RV                 SA,RE,VI,SU,IN,AV          00
PAGE : 1                   EC,VO,SO,RY            00      FI
PAGE 0001-----DERNIERE PAGE 0001
LA NOUVELLE PILE, DECOUVERTE PAR HASARD PAR CL. BIER, PRODUIT DE L'ENERGIE ELEC
TRIQUE A PARTIR D'UN MELANGE HYDROGENE/OXYGENE, EVITANT AINSI LES COMPLICATION
S APORTEES PAR LES DEUX SYSTEMES SEPARES D'APPROXIE D'AIR.
LE FONCTIONNEMENT EST DU A UN NOUVEAU TYPE DE CATALYSE" QUI EST "P
LUS PRES DE LA BIOCHIMIE QUE DE LA CHIMIE CLASSIQUE".

```

II-2. Suppression de chaînes de caractères

Zones variables à renseigner :

- CHAINE 1 : chaîne à supprimer (y compris espaces en tête de chaîne).
- CHAINE 2 : ne pas renseigner.
- BORNE 1 : définissant les bornes d'application.
- BORNE 2 : dans le champ (cf paragraphe Insertion de chaînes de caractères).

Sur notre exemple, je veux supprimer le mot AINSI :

```

MODIF RES.                RF= 700571                PRESSE BASE013
FONCTION STRING SUR LE CHAMP ENTIER:  2
      1 : INSERTION          2 : SUPPRESSION          3 : REMPLACEMENT
      4 : COMPRESSION       5 : FIN
      6 : SUPPRESSION ENTRE 2 BORNES, BORNES COMPRISES
CHAINE-1 : AINSI
CHAINE-2 :
BORNE-1 : EVITANT
BORNE-2 : LES COMPLICATIONS

```

MESSAGE: NB D'OCCURENCES EFFECTUEES PAR CETTE FONCTION: 0000001

Nouveau texte après la suppression :

```

MODIF RES.                RF= 900571                PRESSE BASE013
ORDRE: RV                 SA,RE,VI,SU,IN,??          ??
PAGE : 1                  EC,VC,SV,RV                ST      FI
PAGE 0001-----DERNIERE PAGE :0001
LA NOUVELLE PILE, DECOUVERTE PAR HASARD PAR C. DYER, PRODUIT DE L'ENERGIE ELEC
TRIQUE A PARTIR D'UN MELANGE HYDROGENE/OXYGENE, EVITANT LES COMPLICATIONS APP
ORTEES PAR LES DEUX SYSTEMES SEARES D'ARRIVEE D'AIR.
LE FONCTIONNEMENT EST DU A "UN NOUVEAU TYPE DE CATALYSE" QUI EST "PLUS P
RES DE LA BIOCHIMIE QUE DE LA CHIMIE CLASSIQUE".

```

II-3. Remplacement de chaîne de caractères

Zones variables à renseigner :

- CHAINE 1 : chaîne à remplacer.
- CHAINE 2 : chaîne remplaçante.
- BORNE 1 : définissant les bornes d'application.
- BORNE 2 : dans le champ (cf paragraphe Insertion de chaînes de caractères).

Sur notre exemple, je veux remplacer le mot D'AIR par le mot DE GAZ.

```

MODIF RES.                RF= 900571                PRESSE BASE013
FONCTION STRING SUR LE CHAMP ENTIER:  3
      1 : INSERTION          2 : SUPPRESSION          3 : REMPLACEMENT
      4 : COMPRESSION        5 : FIN
      6 : SUPPRESSION ENTRE 2 BORNES BORNES COMPRISES
CHAINE-1 : D'AIR
CHAINE-2 : DE GAZ
BORNE-1  : D'ARRIVEE
BORNE-2  : .

```

MESSAGE: NB D'OCCURENCES EFFECTUEES PAR CETTE FONCTION : 000001

Nouveau texte après le remplacement :

```

MODIF RES.                RF= 900571                PRESSE BASE013
ORDRE: RV                 SA,RE,VI,SU,IN,.,.,.      ??
PAGE : 1                   EC,VC,SV,RV                ST      FT
PAGE 0001-----DERNIERE PAGE :0001
LA NOUVELLE PILE, DECOUVERTE PAR HASARD PAR C. DYER, PRODUIT DE L'ENERGIE ELEC
TRIQUE A PARTIR D'UN MELANGE HYDROGENE/OXYGENE. EVITANT LES COMPLICATIONS APP
ORTEES PAR LES DEUX SYSTEMES SEPARES D'ARRIVEE DE GAZ.
LE FONCTIONNEMENT EST DU A "UN NOUVEAU TYPE DE CATALYSE" QUI EST "PLUS
PRES DE LA BIOCHIMIE QUE DE LA CHIMIE CLASSIQUE".

```

II-4. Compression de blancs

Il s'agit de remplacer une suite de 'caractères blancs' comprise entre des caractères significatifs, par un seul 'caractère blanc'.

- CHAINE 1 : ne pas renseigner.
- CHAINE 2 : ne pas renseigner.
- BORNE 1 : définissant les bornes d'application.
- BORNE 2 : dans le champ.

Sur notre exemple, je veux éliminer les espaces existants entre la phrase se terminant par DE GAZ et la phrase commençant par LE FONCTIONNEMENT :

```

MODIF RES.                RF= 900571                PRESSE BASE013
FONCTION STRING SUR LE CHAMP ENTIER:  4
      1 : INSERTION      2 : SUPPRESSION      3 : REMPLACEMENT
      4 : COMPRESSION    5 : FIN
      6 : SUPPRESSION ENTRE 2 BORNES, BORNES COMPRISES
CHAINE-1 :
CHAINE-2 :
BORNE-1  : GAZ.
BORNE-2  : LE FONCTIONNEMENT

```

MESSAGE: 'NB D'INCURRENCES EFFECTUEES PAR CETTE FONCTION' = 000001

Nouveau texte après la compression:

```

MODIF RES.                RF= 900571                PRESSE BASE013
ORDRE: RV                 SA,RE,VI,SU,IN,//          ??
PAGE : 1                  EC,VC,SV,RV                ST          FI
PAGE 0001-----DERNIERE PAGE :0001

```

LA NOUVELLE PILE, DECOUVERTE PAR HASARD PAR C. DYER, PRODUIT DE L'ENERGIE ELECTRIQUE A PARTIR D'UN MELANGE HYDROGENE/OXYGENE, EVITANT LES COMPLICATIONS APPORTEES PAR LES DEUX SYSTEMES SEARES D'ARRIVEE DE GAZ. LE FONCTIONNEMENT EST DU A "UN NOUVEAU TYPE DE CATALYSE" QUI EST "PLUS PRES DE LA BIOCHIMIE QUE DE LA CHIMIE CLASSIQUE".

Quand toutes les modifications ont été effectuées, il faut sauvegarder le document final.

On utilise donc la commande (RV). Elle permet de sauver une page que l'on vient de modifier, et de valider le champ modifié.

Le document final est donc :

```

REF      : 900571
TIT.     : ELECTRICITY BY SERENDIPITY
AUT.     : POOL R
SDU.     : SCIENCE
MCL.     : PILE, HYDROGENE, OXYGENE, GAZ
DAT.     : 90/03/92
CLA.     : XXII
CHA.     : INFORMATIONS DIVERSES
DIS.     : 0
RES.     : LA NOUVELLE PILE, DECOUVERTE PAR HASARD PAR C. DYER, PRODUIT DE
          L'ENERGIE ELECTRIQUE A PARTIR D'UN MELANGE HYDROGENE/OXYGENE, EVITANT
          LES COMPLICATIONS APPORTEES PAR LES DEUX SYSTEMES SEARES D'ARRIVEE DE
          GAZ. LE FONCTIONNEMENT EST DU A "UN NOUVEAU TYPE DE CATALYSE" QUI EST
          "PLUS PRES DE LA BIOCHIMIE QUE DE LA CHIMIE CLASSIQUE".

```

On voit donc que cette fonction ST est intéressante pour modifier un champ long tel que les résumés. Mais on ne peut modifier qu'un mot ou groupe de mots après l'autre, on doit toujours revenir au programme initial ST pour effectuer une nouvelle modification.

III. INTRODUCTION DE SYNONYMES DANS UN THESAURUS EXISTANT

J'ai ajouté des synonymes dans 2 bases de données déjà créées qui contenaient toutes les deux déjà un thésaurus.

III-1. Définition d'un thésaurus

Pour MISTRAL, un thésaurus est constitué d'un ensemble de termes sémantiques. Tous les thésauri sont traités de la même façon et utilisent le même langage de commande.

MISTRAL permet de gérer les 3 types de relations de base :

- les relations d'équivalences,
- les relations hiérarchiques,
- les relations associatives.

En effet, MISTRAL gère automatiquement une relation d'équivalence (la synonymie), une relation d'ordre hiérarchique et 10 relations associatives définies par le gestionnaire de la base.

III-2. La relation d'équivalence = la synonymie

Les relations d'équivalence renvoient les divers synonymes d'un concept (en langage naturel) vers le descripteur utilisé dans le langage documentaire.

Tous les documents associés à la classe d'équivalence doivent être retrouvés même si un seul des termes est utilisé comme descripteur.

Dans la relation de synonymie, les termes ont le même sens ou presque et sont pratiquement interchangeable. La relation de synonymie est une relation d'équivalence au sens mathématique du terme c'est à dire que si A est synonyme de B, alors B est synonyme de A.

MISTRAL gère un fichier Inverse de Synonymie (FISY) dans lequel on trouve pour chaque classe de synonymes, un article regroupant la liste de tous les documents indexés par l'ensemble des termes de la classe; ce qui améliore le temps de réponse pour les recherches mettant en jeu ce type de relations. En effet, ce procédé évite de fusionner les listes des documents indexés pour chacun des termes de la classe.

On indique une relation de synonymie en écrivant les 2 termes synonymes séparés par un signe égal (=).

Il est possible d'avoir plusieurs termes synonymes.

Ex : FIRME = SOCIETE = COMPAGNIE

Des espaces ou des blancs sont autorisés entre les mots-clés et le signe égal.

L'ensemble des mots-clés cités dans la commande va constituer une classe de synonymes.

Quand plusieurs termes sont synonymes, on peut choisir parmi eux un synonyme préférentiel. C'est ce terme qui sera édité de préférence aux autres dans une édition du thésaurus. Il faut alors mettre deux '**' devant le terme qui sera désigné comme synonyme préférentiel.

Ex : ** CONGRES = COLLOQUE.

III-3. Programme qui permet d'ajouter les synonymes au thésaurus existant

Je n'utiliserai pas le cas de synonyme préférentiel pour les 2 bases où je vais ajouter les synonymes. J'utiliserai donc uniquement le signe égal (=).

III-3-1. Organisation générale

La chaîne de gestion des thésauri se réalise en 3 étapes successives :

- *Première étape*

Préparation de la gestion des relations. Suivant le type de relations, une des 2 phases suivantes est nécessaire :

- * Programme FRTHEs
pour les relations standard et optionnelles; les relations entre documents ne sont introduites que dans cette phase;
- * Programme FRTAUT
pour les relations des mots d'un groupe de mots.

- *Deuxième étape*

- * Programme MJTHS2
Gestion des lexiques et thésauri (fichiers Liens, Relations, FIS et FID).

- *Troisième étape*

- * Programme MJTHS3
Gestion de la classe des Synonymes (FIS). La synonymie n'ayant aucune signification pour les documents, cette étape ne concerne que les descripteurs.

Je n'expliquerai que le programme FRTHEs qui permet d'introduire les synonymes dans le thésaurus existant. Les autres programmes sont faits par l'opérateur de MISTRAL.

III-3-2. Le programme FRTHEs

Les données à entrer dans l'ordre sont :

- article mot de passe,
- articles paramètres,
- commandes.

* article mot de passe

IDEN MOT-PASSE

MOT-PASSE = mot de passe pour les chaînes batch entré dans ADMIST.

* articles paramètres

$$\left[\text{PARM} \left[\text{IN} : \begin{cases} \text{CARD} \\ \text{FILE} \end{cases} \right] \left[, \text{KWL} : 1 \right] , \left[\text{SEGN} : \begin{cases} \text{YES} \\ \text{NO} \end{cases} \right] 1, \begin{cases} \text{LIST} \\ \text{NLST} \end{cases} \left[, \text{NPKW} \right] \right]$$

- IN : CARD : entrée des commandes depuis le même membre de librairie que les articles mot de passe et paramètres.
- IN : FILE : entrée des commandes depuis un support magnétique.
- KWL : longueur maximale des descripteurs présents dans les commandes; par défaut : valeur donnée à l'initialisation de la base ADMIST.
- SEGN : sigle contrôle de séquence (entrée CARD seulement).
- YES : sigle oui.
- NO : sigle non.
- LIST : sigle édition des commandes.
- NLST : sigle de non édition des commandes (implicite).
- NPKW : sigle de non édition des descripteurs; si ce paramètre est absent, il y a édition des descripteurs en ordre alphabétique, suivis des numéros de commandes où ils apparaissent.

* commandes : désignation du champ

FLD < indicatif du champ >

FLD est le sigle du champ.

* L'introduction de synonymie se présente sous forme d'une suite de descripteurs reliés par l'opérateur de synonymie (=).

Le programme pour entrer les synonymes est donné ci-dessous :

```

S : LMN SL .V5DON
C : FSE
F : CREATE B13-FRTHES
  IDEN PRESSE
  PARM IN:CARD, LIST
  FLD MCL.
  CAO = CONCEPTION ASSISTEE ORDINATEUR
                                     entrée
                                     des
                                     synonymes

/                                     <-- fin de traitement

```

III-3-3. Etape de vérification

Je vais vérifier que les synonymes que j'ai entré sont bien insérés au thésaurus de la base.

a) Interrogation sur le thésaurus

Cette interrogation se fait par la commande :

```
M : TH[ESAURUS] terme
```

terme : mot ou groupe de mots appartenant au lexique.

Cette commande permet d'obtenir la liste des termes reliés à un terme donné.

Je veux donc rechercher si le terme CAO à un ou des synonymes. Je tape donc :

```
M : TH CAO
```

```
?m:th cao
```

```

                RES. NIVEAU
          T1          3          CAO
SYN T2          1          CONCEPTION ASSISTEE ORDINATEUR
Selectionner ou non ?

```

Cette commande me donne donc :

- tous les synonymes associés au terme demandé. Ici, il n'y en a qu'un qui est CONCEPTION ASSISTEE ORDINATEUR.
- le nombre de documents indexés par le terme et ses synonymes (RES). Ici, on a un document qui contient le sigle développé et 3 documents qui contiennent le sigle CAO ou le sigle développé. T1 = premier élément de la liste initiale.

Pour visualiser les documents répondant à cette commande, il suffit de répondre à la question par SELECTIONNER.

b) Interrogation sur le lexique

Cette interrogation se fait par la commande :

M : LE(XIQUE) terme

terme = terme cible à partir duquel s'effectue l'édition.

Cette commande permet de voir la liste des termes alphabétiquement voisins d'un terme donné.

En tapant donc :

M : LE CAO

?m:le cao

	RES.	REL.	
T1	4		CANON GAZ LEGER
T2	1		CANON INDUCTION
T3	5		CANON NUCLEAIRE
T4	2		CANON PLASMA
T5	3		CANON RAIL
T6	4		CANDPUS
T7	3	*	CAO
T8	3		CADOTCHOUC
T9	9		CAPACITE
T10	10		CAPELLA
T11	2		CAPOT
T12	4		CAPRI
T13	1		CAPRICORNE
T14	11	*	CAPTEUR
T15	25		CAPTEUR OPTIQUE
T16	2		CAPTEUR PRESSION
T17	4	*	CAPTURE
T18	18		CAPTURE NEUTRONIQUE

Monter (M) x/descendre (D) x/Selectionner/Non ?
?

T7 : numéro d'ordre à la liste affichée sur l'écran.

RES : nombre de documents indexés par le terme et ses synonymes.

REL : existence de relation(s) associée(s) à chaque terme (signalée par un astérisque *).

En bas de l'écran s'affiche un message qui est :
MONTER (M) x/Descendre (D) x/Selectionner/ Non?

M : pour poursuivre l'édition des n termes
alphabétiquement inférieurs.

D : pour poursuivre l'édition des n termes
alphabétiquement supérieurs.

S : pour sélectionner tous les termes de la liste.

J'ai effectué la même commande mais en utilisant le sigle
développé CONCEPTION ASSISTEE ORDINATEUR pour vérifier que
l'on obtenait le même nombre de réponses.

?m:le conception assistee ordinateur

	RES.	REL.	
T1	13		COMPTE-RENDU
T2	3	*	COMPTEUR
T3	8		COMPTEUR PROPORTIONNEL
T4	1	*	COMPTON
T5	16	*	CONCENTRATION
T6	18	*	CONCEPTION
T7	3	*	CONCEPTION ASSISTEE ORDINATEUR
T8	1		CONCEPTION ASSISTEE ORDINATEUR
T9	8		CONDENSATEUR
T10	8	*	CONDITION
T11	2		CONDITION METEOROLOGIQUE
T12	1		CONDITIONNEMENT
T13	3	*	CONDUCTEUR
T14	1	*	CONDUCTIBILITE
T15	12		CONDUCTIBILITE ELECTRIQUE
T16	1		CONDUCTIBILITE THERMIQUE
T17	1	*	CONDUCTION
T18	2		CONDUCTION THERMIQUE

Monter (M) x/descendre (D) x/Selectionner/Non ?

J'ai obtenu le même nombre de réponses, je peux donc conclure
que les synonymes que j'ai entrés sont bien incorporés au
thésaurus.

CONCLUSION

Comme le montre ce rapport, MISTRAL est un logiciel puissant, complet de traitement de l'information.

En effet, il permet :

- saisie, indexation automatique
- recherche, interrogation
- élaboration de thésaurus
- gestion de la base (lexiques, thésaurus...)

De plus MISTRAL sur un centre serveur permet la consultation simultanée de 25 bases.

Enfin, il offre de nombreuses facilités pour protéger l'accès aux bases documentaires en consultation :

- identification de l'utilisateur
- liste des usagers habilités
- mot de passe
- clé d'habilitation.

De même, pour obtenir de bons résultats lors d'interrogation, il convient de :

- définir des règles d'écriture et de les suivre
- bien remplir les champs
- faire une bonne indexation.

Il en résulte donc que, la mise en place d'une base de données doit être définie avec précision et rigueur car elle sera ensuite délicate à modifier; ce qui nécessite un contact permanent avec le documentaliste qui demande la création de la base et l'administrateur qui effectue sa création.

ANNEXE 1

ORDRE ET SIGNIFICATION DES SIGLES DES PROCEDURES

Le sigle d'une procédure est composé de 2 caractères. Ces sigles sont tous différents.

QU (ESTION)	:	Recherche
LE (XIQUE)	:	Affichage du lexique
TH (ESAUROS)	:	Affichage du thesaurus
DA (TE)	:	Intervalle de Mise à jour
EX (ECUTER)	:	Réexécution d'une stratégie de recherche, exécution d'une macro-procédure
VI (SUALISER)	:	Edition des documents sur le terminal
ED (ITER)	:	Edition des documents en différé
HI (STORIQUE)	:	Edition des instructions de recherche
IN (DEX)	:	Edition des références entre instructions de recherche
EF (FACER)	:	Effacement des instructions de recherche
NU (MEROTER)	:	Rénumérotation des instructions de recherche
EN (VIRONNEMENT)	:	Edition du contexte de la recherche
CH (AMP)	:	Edition des informations sur les champs
RE (LATION)	:	Liste des relations existant dans la base
LA (NGUE)	:	Liste des langues disponibles pour l'édition des lexiques.
DO (MAINE)	:	Définition des domaines interrogés
DE (MANDEUR)	:	Identification de la personne pour qui est effectuée la recherche

BA (SE)	:	Définition de la base interrogée
CO (MPTABILITE)	:	Comptabilité
ZE (RO)	:	Remise à zéro du compteur temps relatif
HE (URE)	:	Edition de la date et de l'heure
HO (RLOGE)	:	Edition du temps passé
OP (TION)	:	Définition des options de traitement
FO (RMAT)	:	Définition des paramètres de composition de texte
FI (N)	:	Déclaration de fin de session
ME (NU)	:	Liste des procédures
SO (S)	:	Assistance sur une procédure
AD (RESSE)	:	Adresse de routage des éditions
TE (LEX)	:	Réception de message
AT (TENTE)	:	Arrêt momentané de la session
BI (BLIOTHECAIRE)	:	Gestion du fichier Bibliothécaire
MA (CRO)	:	Exécution des macro-procédures
OU (TPUT)	:	Sortie magnétique
SA (ISIE)	:	Saisie, mise à jour des documents
PR (OFIL)	:	Exécution des profils
RA (PPORT)	:	Exécution des rapports
VR (OOM)	:	Recherche express
PE (RIODE)	:	Modification période des domaines
XX	:	Inutilisé dans cette version
AR (BRE)	:	Visualisation arborescence
ZO (OM)	:	Visualisation des documents et fenêtres de pertinence
PA (GINATION)	:	Pagination ou visualisation page par page
FM (FORMULAIRE)	:	Interrogation de MISTRAL à l'aide de formulaires, c'est-à-dire d'écrans simples
TW (RITER)	:	Envoi du contenu des documents sur imprimante

Sigles des commandes

Il est impératif que les sigles des commandes soient différents les uns des autres.

TX (TE)	:	commande de recherche sur texte
TR (I)	:	commande de tri pour l'édition en différé
TI (TRE)	:	commande de titre pour l'édition en différé
FI (N)	:	fin d'analyse des commandes pour l'édition en différé
RF	:	commande de recherche sur référence
QI	:	question sur intervalle de valeurs (réservé pour extension future)
SD	:	sélection déduite

Sigle des réponses

O (UI)	:	Réponse affirmative
N (ON)	:	Réponse négative
M (ONTER)	:	Edition des termes alphabétiquement inférieurs
D (ESCENDRE)	:	Edition des termes alphabétiquement supérieurs
S (ELECTION)	:	Sélection des termes
Z1 à Z5	:	(réservés par extension future)
A (VEC)	:	
F (ICHE)	:	Demande de visualisation microfiche
V (IDEO)	:	Demande de visualisation vidéo disque
C (ONTINUER)	:	Continuer la visualisation
E (LIMINATION)	:	Elimination du document
Q (UIT)	:	Sortie éditeur
I (NSERT)	:	Insertion ligne
L (IST)	:	Affichage ligne
P (RINT)	:	Affichage ligne avec Numérotation
D (ELETE)	:	Suppression ligne ou caractère
R (EPLACE)	:	Remplacement ligne ou caractère

C (OPY)	:	Copie ligne
T (TRANSFERT)	:	Ligne
V (ERIFY)	:	Mode Vérification
H (ELP)	:	

Caractères de fonction

Les caractères doivent être différents les uns des autres.

(:	Parenthèse gauche
)	:	Parenthèse droite
#	:	Masque
?	:	Troncature limitée
+	:	Troncature non limitée
'	:	Séparateur servant à encadrer les mots ou caractères réservés
:	:	Préfixe Indicatif de Champ
,	:	Séparateur de deux opérandes pour recherche sur texte
;	:	Fin de commande
-	:	Caractère de continuation (Suite)
/	:	Slash
*	:	Séparateur de paramètres implicites dans une macro-procédure



Sigles des opérateurs

4 caractères peuvent être pris en compte, par opérateur.

Les sigles doivent être différents les uns des autres.

ET	:	Opérateur 'ET'
OU	:	Opérateur 'OU'
SAUF	:	Opérateur 'SAUF'
NON	:	Opérateur de négation
DST	:	Opérateur de distance

ADJ	:	Opérateur d'adjacence
PH	:	Phrase
PA	:	Paragraphe
RCH	:	Opérateur de recherche de chaîne de caractères
CHP	:	Champ
DATE	:	Date
#	:	Opérateur "Différent de"
<	:	Opérateur "inférieur"
>	:	Opérateur "supérieur"
=	:	Opérateur "Egal"
>=	:	Opérateur "Plus grand ou Egal"
<=	:	Opérateur "Plus petit ou Egal"
POUR	:	Documents obtenus par intersection de deux listes résultat
PAR	:	Idem POUR, mais union au lieu d'intersection
ET#	:	Opérateur booléen sur paragraphe
ET.	:	Opérateur booléen sur phrase
SF#	:	Opérateur booléen sur paragraphe
SF.	:	Opérateur booléen sur phrase
PARM	:	Directive introduisant les paramètres formels d'une macroprocédure
PROC	:	Directive introduisant les procédures dans les macroprocédures

ANNEXE 2

B13_MOTVID
10 IDER NR
20 PARM INIT
30 FLD TIT.
40 A
1 AFTER
2 AINSI
3 ALL
50 AN
60 AND
1 ANNUAL
2 APRES
3 ARE
70 AS
80 AT
90 AU
1 AUSSI
100 AUX
1 AVANT
2 AVEC
110 B
1 BE
2 BEGIN

+++

3 BEGINS
4 BUT
5 BY
120 C
130 CA
1 CAN
2 CAS
150 CE
1 CES
160 CET
170 CETTE
180 CHEZ
1 COMME
2 COULD
220 D
230 DA
240 DANS
250 DAS
260 DE
270 DEM
280 DEN
290 DER
300 DES

+++

310 DIE
320 DU
330 E
340 EN
 1 ENCORE
 2 ENTRE
 3 EST
350 ET
 1 ETRE
 2 EXAMINES
360 F
 1 FOR
370 FROM
380 G
390 H
 1 HAS
 2 HAVE
 3 HOW
400 I
 1 ICI
 2 IL
410 IN
 1 INTO

+++

 2 IS
420 IT
 1 ITS
430 J
440 K
450 L
460 LA
470 LE
480 LES
 1 LEUR
 2 LEURS
 3 LORS
490 M
 1 MAIS
500 ME
 1 MEME
 2 MEMES
 3 MIEUX
510 N
520 NE
 1 NEW
 2 NEXT
 3 NI

+++

4 NO
5 NON
6 NOT
7 NOTRE
530 O
540 OF
550 ON
1 ONT
2 OTHER
560 OU
1 OVER
570 P
580 PAR
1 PAS
2 PENDANT
3 PEUT
4 PLUS
5 FOUR
6 PRES
590 Q
1 QU
600 QUE
1 QUI

+++

610 R
620 S
630 SA
1 SANS
2 SAY
3 SAYS
640 SE
1 SEIN
2 SELON
650 SES
1 SI
2 SON
3 SONT
660 SOUS
670 SUR
680 T
690 TE
700 THAN
710 THAT
720 THE
1 THEIR
730 THEM
740 THEN

+++

750 THIS
760 TO
1 TOO
2 TOUS
3 TOUT
4 TOUTE
5 TOUTES
6 TRAVERS
770 UN
1 UNDER
780 UNE
1 UP
2 USED
790 V
1 VERS
800 W
1 WHAT
2 WHY
3 WILL
4 WITH
5 WOULD
810 Y
820 Z

+++

1 -
2 '

C:



BIBLIOTHEQUE DE L'ENSSIB



8017118