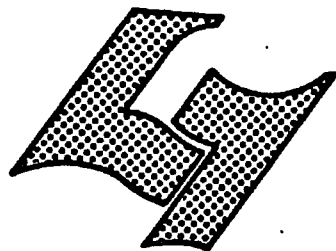


UNIVERSITE CLAUDE BERNARD LYON-I  
43, Boulevard du 11 Novembre 1918  
69621 VILLEURBANNE

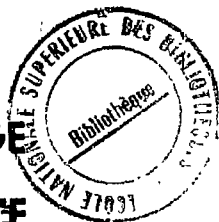


## *Diplôme d'Etudes Supérieures Spécialisées*

# informatique documentaire

\* MEMOIRE DE STAGE

\* ~~NOTE DE SYNTHESE~~



IMPLANTATION D'UNE BASE DE DONNEES SUR  
UN SERVEUR .

AUTEUR : G A U S S O T Catherine

DATE : Juin 1979

DESS  
1979  
17

## R E M E R C I E M E N T S

A l'issue de ce stage de quatre mois, je voudrais remercier la SPI d'avoir accepté de m'accueillir, et plus particulièrement, le service SPIDEL .

Que Madame STERN et Monsieur GENTRIC, qui m'ont conseillée et guidée, trouvent ici l'expression de ma gratitude, à laquelle ont droit également l'équipe SPIDEL et le service technique que j'ai souvent mis à contribution .

## S O M M A I R E

Remerciements .

Sommaire .

Introduction . . . . . p. 1

1. Pourquoi et comment choisir une base de

données . . . . . p. 3

1.1. Etude de marché . . . . . p. 3

1.2. Etude des besoins des utilisateurs . . . . . p. 8

1.3. La recherche des bases . . . . . p. 11

1.4. Analyse du contenu d'une base . . . . . p. 11

1.5. Coût de la recherche d'une base . . . . . p. 12

2. Définition des besoins pour implanter des bases

de données . . . . . p. 13

2.1. La croissance de l'information . . . . . p. 13

2.2. Les moyens . . . . . p. 13

3. Implantation d'une base dans STAIRS . . . . . p. 16

3.1. Définition d'une base dans STAIRS et procé-  
dure d'entrée des documents . . . . . p. 16

3.2. Contrainte STAIRS . . . . . p. 24

3.3. Les caractères spéciaux . . . . . p. 25

3.4. Le secret . . . . . p. 27

3.5. Chargement et contrôle d'une base . . . . . p. 33

3.6. Mises à jour . La sauvegarde des fichiers . . p. 34

4. Gestion d'un serveur . . . . . p. 39

4.1. Le fichier adresse . . . . . p. 39

4.2. La documentation payante . . . . . p. 40

4.3. La facturation . . . . . p. 40

Conclusion . . . . . p. 41

Annexes

## I N T R O D U C T I O N

La SPI, Société Pour l'Informatique, filiale de Pechiney Uguine Kuhlman, a décidé, en collaboration avec Madame Stern, de lancer une opération "gros serveur" sur le marché français .

Le service SPIDEL, Service pour l'Information et la Documentation en Ligne, est un service de consultation de bases de données en conversationnel par le logiciel STAIRS ; ce service, opérationnel depuis février 1979, a été présenté officiellement à la presse et au monde de la documentation le 15 mai 1979 au cours d'une conférence de presse organisée par la SPI .

SPIDEL veut répondre aux besoins des entreprises et met ainsi à leur disposition des informations couvrant surtout les domaines techniques et économiques. Pour l'instant, les bases de données en ligne sont en français et traitent d'abord des problèmes nationaux et européens . La SPI participe à l'essor européen de la documentation en étant serveur sur EURONET .

Le sujet du rapport de stage, "Implantation d'une base de données sur un serveur", a été choisi en accord avec Monsieur Bouché, bien que mon travail pendant quatre mois ait porté sur un autre domaine, parce que le sujet était original . En effet, aucun stage ne s'est encore déroulé chez un serveur .

Mon travail quotidien était de faire des interrogations, des démonstrations dans des manifestations comme A.D.B.S., les J.I.I.A., INOVA, d'assurer une partie des relations publiques en répondant aux appels téléphoniques des clients ou des personnes intéressées . J'ai également dû préparer des pochettes de documentation com-

merciale SPIDEL .

Le rapport s'articule de la façon suivante :

1. Pourquoi et comment choisir une base de données ?
2. Définition des besoins pour implanter une base de données .
3. Implantation d'une base dans STAIRS .
4. Gestion .

Je devais, outre l'aspect technique, élaborer une sorte de cahier des charges de l'implantation d'une base, mais je n'ai pas pu obtenir les coûts exacts des différentes opérations . En effet, les lois de la concurrence ne permettent pas la divulgation des prix de revient exacts .

1. POURQUOI ET COMMENT CHOISIR UNE BASE  
DE DONNEES.

## 1. POURQUOI ET COMMENT CHOISIR UNE BASE DE DONNEES .

### 1. 1. Etude de marché . \*

1. 1. 1. L'information vise un marché potentiel articulé en quatre groupes très distincts qui demandent une approche différente .

#### 1. 1. 1. 1. Les "grandes" sociétés

sont les 1000 à 1500 premières sociétés françaises déjà familiarisées avec l'exploitation de l'information à tous les niveaux .

Elles ont un service de documentation important et structuré . L'information circule bien entre elles et elles sont souvent membres d'organisations professionnelles comme l'A.D.B.S.

Elles sont déjà équipées de terminaux d'interrogation et ont l'habitude de se connecter à d'autres systèmes documentaires informatisés . Ce sont donc des sociétés très réceptives et sensibles à de nouvelles propositions d'information, mais, habituées à des services élaborés, elles seront exigeantes sur la qualité du service offert .

#### 1. 1. 1. 2. Les sociétés "moyennes"

dont le nombre s'évalue en une dizaine de milliers ressentent des besoins d'information comparables aux "grandes", mais n'ont en général pas les moyens d'entretenir un service de documentation très important . Si ce dernier existe, il ne possède en général pas de terminal d'interrogation .

#### 1. 1. 1. 3. Les "petites" sociétés

n'ont, le plus souvent, pas de service de documentation à cause de leur petite taille ou parce qu'elles n'en ressentent pas le besoin . Cette situation est très fréquente dans la majorité des entreprises .

#### 1. 1. 1. 4. Les intermédiaires

constituent une catégorie originale et très importante . Ils sont des professionnels de l'information et touchent un très vaste public auquel ils offrent un service complet d'analyse et de synthèse des réponses aux questions posées par exemple .

\* Note interne manuscrite de M. BRESSANGE : Projet de plan de lancement de l'opération "Gros Serveur".

Les courtiers en information, comme le bureau Marcel VAN DIJK, les chambres de commerce, les ARIST, les observatoires économiques de l'INSEE, et les fournisseurs de bases eux-mêmes, sont donc des prescripteurs intéressants et de bons clients .

Il ne faut pas non plus négliger le marché international, et les pays de langue française qui forment un public privilégié, car les informations fournies par SPIDEL sont en priorité en français .

Le démarrage d'EURENET et la participation de la SPI en tant que serveur devraient aider à la commercialisation du produit .

1. 1. 2. Evaluation du marché potentiel constitué par les différents secteurs d'activité .

1. 1. 2. 1. Une étude de la répartition des membres de l'A.D.B.S. dans les 500 premières sociétés françaises, puis dans les 5000 premières (d'après une classification du "Nouvel Economiste") fournit une indication précieuse sur les secteurs d'activité particulièrement concernés par les problèmes documentaires, donc plus prêts à accepter l'activité SPIDEL .

Les tableaux des pages suivantes classent ces secteurs en quatre catégories :



**INTERET DES SECTEURS D'ACTIVITE POUR  
LA DOCUMENTATION \***

SECTEURS D'ACTIVITE	Nb. de Sociétés	
	dans les 500 premières	dans les 5000 premières
<b>1 - Secteurs très concernés</b>		
Chimie lourde	17	81
Ingénierie	6	147
Matériel et construction électrique	36	222
Métaux non ferreux	4	43
Parachimie	18	122
Pétrole, Gaz, Carburant	15	51
Publicité	4	106
Transport	18	129
Edition, Presse, Imprimerie	5	74
	<hr/> 123	<hr/> 975
<b>2 - Secteurs intéressés</b>		
Aéronautique	28	140
Divers	8	35
Extraction	2	32
Industries alimentaires	40	247
Mécanique	26	255
Mécanique de précision	6	75
Papiers, Cartons, Emballages	8	92
Auxiliaires de transport	2	48
Service Public	17	57
Sidérurgie, Metallurgie	27	217
Sociétés de commerce	6	286
Tourisme, Hôtellerie, Loisir	12	72
Travaux Publics	15	121
	<hr/> 197	<hr/> 1679

\* BRESSANGE (A.)

Note interne manuscrite de M. BRESSANGE : Etude du marché SPIDEL, octobre 1978 .

**INTERET DES SECTEURS D'ACTIVITE POUR LA DOCUMENTATION (suite)**

SECTEURS D'ACTIVITE	Nb. de Sociétés	
	dans les 500 pre- mières .	dans les 5000 pre- mières
<b>3 - <u>Secteurs peu concernés</u></b>		
Automobile, Cycle	23	103
Bâtiment	10	261
Reissen	20	94
Immobilier	31	50
Location de matériel, Leasing,		
Travail temp.	14	68
Matériaux de construction	11	115
Textile	11	137
Verre , Céramique, Faïence	3	49
	<hr/> 123	<hr/> 877
<b>4 - <u>Secteurs pas concernés</u></b>		
Caoutchouc	2	31
Chaine de magasins	22	96
Construction et réparation		
navale	3	17
Coopérative de consommation	5	10
Cuir et chaussures	2	45
Grands magasins	9	42
Habillement	1	59
Import - Export	8	65
Industrie du bois	1	76
Industrie des plastiques	3	76
	<hr/> 57	<hr/> 517
<b><u>TOTAL</u></b>	<b>500</b>	<b>4048</b>

Les secteurs de la chimie, l'ingénierie, la construction électrique et la mécanique ont des besoins particulièrement marqués en information.

1. 1. 2. 2. Les organismes suivants sont aussi membres de l'A.D.B.S.

TYPE D'ORGANISME *	Nb. de Membres de l'A.D.B.S.
Laboratoire, Centre de recherche, bureau d'études	162
Lycée, Ecole, Faculté, Université, IUT	71
Association, Union, Syndicat, Organisa- tion professionnelle	63
Bibliothèques, Archives, Centre d'infor- mation	60
Administration	35
Ministère	22
Organisme para public	15
Préfecture	6
Mairie	5
	<hr/> 441

\* BRESSANGE (A.)

1. 1. 3. Conclusions à tirer de cette étude.

Le marché français des interrogations en ligne est passé de 14 à 18 millions de francs en 1978, le marché européen de 68 à 88 millions de francs, ce qui représente une augmentation sensible, respectivement de 28,5% et de 29,5%.

Ce marché a connu une croissance moyenne de 130% par an de 1971 à 1977 qui devrait rester supérieure à 40% par an de 1978 à 1982. Mais il faut observer l'apparition de nouveaux domaines d'intérêt. En effet le marché actuel, centré sur la chimie et la médecine est proche de la saturation, tandis que la demande dans les secteurs juridiques et économiques, longtemps délaissée, est en plein essor.

La SPI a donc une occasion de pénétrer le marché en proposant des bases originales dans des créneaux nouveaux : économiques, juridiques, sociaux .... Mais elle ne devra pas concurrencer, dans un premier temps, des gros serveurs déjà bien implantés; en effet, les cinq plus gros serveurs occupent 93% du marché.

## 1. 2. Etude des besoins des utilisateurs .

### 1. 2. 1. Etude générale.

Une étude de besoins dans les chambres de commerce et de l'industrie confirme l'existence de créneaux intéressants . Cette étude a été faite d'une façon empirique : les enquêteurs se sont installés dans les diverses chambres de commerce et ont noté les questions posées par les entreprises . Un travail de synthèse fait apparaître que les questions ponctuelles sont beaucoup plus fréquentes (70%) que les questions de type documentaire (30%) .

Les questions se répartissent dans les différents domaines suivants :

- économique 14%

(80% de ces questions concernant l'économie régionale)

- juridique 12 à 13%

- social 10%

- sur la formation 10%

- technique, technologique 10% ....

Etant donné que 70% des questions sont ponctuelles, il est difficile de trouver des bases de données informatisées. Les 30% de demandes bibliographiques créent une difficulté car la majorité des questions sont d'ordre économique . Or la durée de l'information économique est très courte . Elle ne dépasse que rarement les six mois . Il est intéressant de remarquer que le domaine scientifique, où l'information est la plus abondante, n'intéresse que très peu l'entreprise. Sur l'ensemble des entreprises françaises, l'information scientifique ne représente même pas 15% de l'ensemble des besoins documentaires .

### 1. 2. 2. Un exemple précis : la chambre de commerce et d'industrie de Lyon :

La chambre de commerce de Lyon, à l'instar des autres chambres, est un établissement semi-public, semi-privé, visant la meilleure progression de l'économie dans sa circonscription . Elle est à la disposition de toutes les entreprises de sa région . L'information dispensée est souvent plus utile aux petites et moyennes entreprises qu'aux grandes, dotées d'un service de documentation . La chambre est à l'écoute des souhaits de l'entreprise .

D'après l'étude statistique sur les questions posées, il ressort que les créneaux que le centre de documentation doit considérer comme prioritaires et sur lesquels il doit s'attacher à être exhaustif sont :

- la vie et la gestion de l'entreprise : management, méthode de gestion, marketing, ...
- les études sectorielles de branches industrielles : avant tout au niveau régional, mais également, dans la mesure du possible, au niveau national .
- dans une moindre mesure, l'aménagement régional, en liaison avec des organismes extérieurs comme la mission régionale ou la DATAR .

#### Demandes d'articles en 1976, classées par thèmes

##### Vie et gestion de l'entreprise

- relations sociales dans l'entreprise	10,5 %
- formation professionnelle continue	6,6 %
- condition de travail dans l'entreprise	9,2 %
- administration générale et direction de l'entreprise	12,2 %
- législation-fiscalité	3,4 %
- autres aspects de la gestion de l'entreprise	5 %
- application de l'informatique à la gestion de l'entreprise	2,3 %
- innovation, recherche	3,2 %
- études par secteurs industriels ou par taille d'entreprise (multinat. , PMI)	6,2 %
	<hr/>
	58,6 %

Aménagement du territoire .

- Régionalisation	11,8 %
- Urbanisme, rénovation urbaine	3 %
- Transports	3,5 %
- Tourisme-hôtellerie	1,5 %
	<hr/>
	19,8 %

Questions socio-économiques .

- Problèmes catégoriels de travailleurs	3,6 %
- Emploi-chômage	3,4 %
- Salaires, retraites, assurances sociales	5 %
- Distribution, urbanisme commercial	3,4 %
- Economie de pays étrangers	2,3 %
- Conjoncture économique	2,7 %
- Divers	0,8 %
	<hr/>
	21,2 %

Total : # 100 %

### 1. 3. La recherche des bases .

La prospection se fait dans tous les organismes qui oréent les bases et les banques, mais en s'aperçoit que la situation est assez médiocre . En effet, la plupart des bases sont manuelles .

Une base une fois trouvée, en commence l'étude de son contenu en fonction des créneaux déterminés .

Par exemple, pour le CETIM, d'après toutes les questions posées par les entreprises, la fréquence des questions sur la mécanique était très élevée en pourcentage . Or y avait-il sur le marché d'autres services susceptibles de satisfaire ces besoins ? La réponse étant négative, il fallait étudier les bases proposées par des organismes de la mécanique .

Si la base trouvée convient, en étudie son contenu . La SPI a choisi jusqu'à présent des bases informatisées .

### 1. 4. Analyse du contenu d'une base .

Le service SPIDEL a adopté une solution systématique pour rédiger la notice descriptive d'une base de données . Le plan reste le même d'une base à l'autre afin que l'utilisateur puisse toujours trouver d'une façon rationnelle les informations qu'il attend : (voir annexe 1.4.)

- le fournisseur de la base
- analyse du contenu des documents
- domaines couverts par la base
- description d'un document
- approche d'une base
- secret
- accès aux documents primaires ...

Le document peut être élaboré, soit en compagnie du fournisseur - dans ce cas sa rédaction est rapide - soit d'après des documents procurés par ce dernier . La procédure est alors un peu plus longue ; en effet une première version est envoyée au fournisseur pour des modifications

éventuelles . Celui-ci doit donner son accord pour la version définitive qui peut n'être adoptée qu'après un certain nombre de navettes . Ainsi le temps passé à cette étude peut varier grandement : d'une journée effective à plusieurs; les délais sont aussi très lâches .

#### 1. 5. Coût de la recherche d'une base .

Il m'a été difficile de calculer un coût global pour la recherche d'une base . Aussi ai-je laissé les résultats bruts .

##### - Combien de temps pour trouver une base ?

1 semaine d'une personne de la direction .

##### - Combien de temps pour parvenir à la signature du contrat ?

2 à 3 mois de délai

1 semaine de travail effectif de quelqu'un de la direction .

##### - Combien de temps pour analyser le contenu d'une base ?

2 jours d'une personne de la direction .

##### - Frais généraux : déplacements, divers ...

environ 10000 francs .



2. DEFINITION DES BESOINS POUR IMPLANTER

DES BASES DE DONNEES .

## 2. DEFINITION DES BESOINS POUR IMPLANTER DES BASES DE DONNEES.

### 2. 1. La croissance de l'information .

La croissance exponentielle de l'information pose de gros problèmes . En effet, devant le volume des informations, les capacités de stockage traditionnelles ne peuvent suffire, et ne permettent pas de consulter rapidement le document recherché .

Trois besoins fondamentaux apparaissent donc :

- le stockage d'une grande quantité d'informations .
- l'accès rapide à une information précise .
- la nécessité d'un langage documentaire évolué .

De telles exigences impliquent des moyens informatiques très sophistiqués et très performants .

### 2. 2. Les moyens .

#### 2. 2. 1. Les capacités de traitement de la SPI et le réseau de distribution .

Le centre de calcul dispose de deux ordinateurs couplés :

- un IBM 30-32 de 6 millions de caractères de mémoire centrale (6000 K)
- un IBM 158 de 5 millions de caractères de mémoire centrale (5000 K) .

Sous peu, l'IBM 158 sera remplacé par un autre IBM 30-32, mais loué, en attendant l'achat de la série H IBM qui sortira en 1980 et qui disposera d'une capacité mémoire presque égale à deux fois celle d'un 30-32 .

Les deux unités centrales peuvent se dépanner mutuellement en cas d'indident .

Un nombre élevé de périphériques très élaborés leur sont connectés . Il faut citer :

- des unités de disques d'une capacité de 16 milliards de caractères en ligne .

- 18 dérouleurs de bandes .
- 5 imprimantes .
- 2 contrôleurs de transmissiens.

Les périphériques sont reliés simultanément aux deux ordinateurs .

La SPI utilise tous les moyens de transmission disponibles en France :

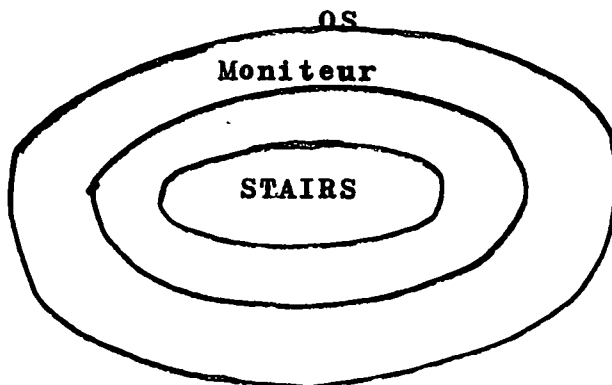
- des lignes louées .
- le réseau téléphonique commuté .
- le telex ...
- les réseaux .
- le réseau TRANSPAC (SPI, par l'intermédiaire de PUK, est membre fondateur d'UTIPAC, la société des utilisateurs du réseau TRANSPAC) .

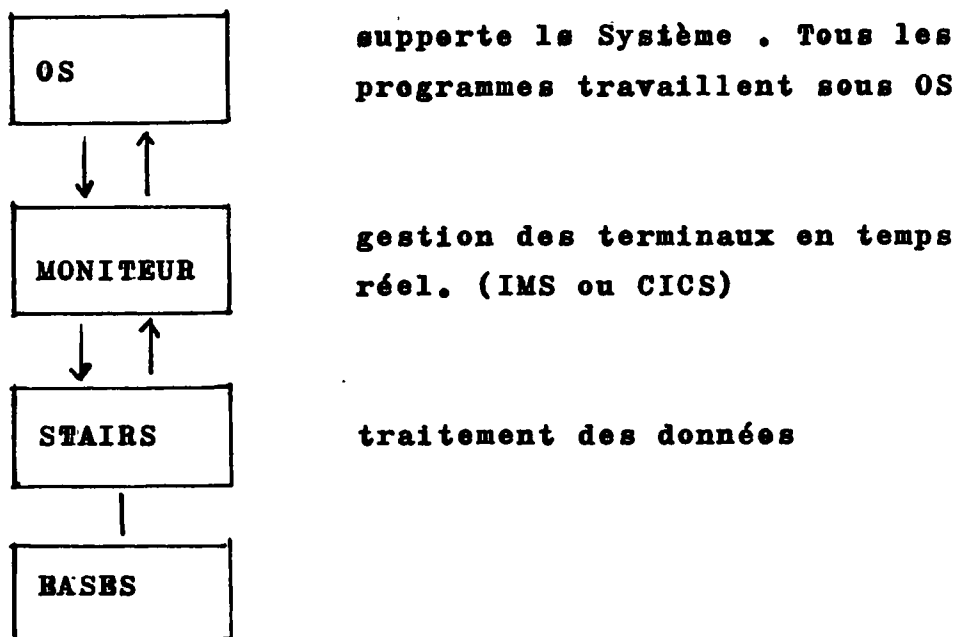
La possibilité d'adaptation à un réseau, la mémoire à accès direct sur disque, la puissance de calcul, les possibilités de traitement, et les périphériques, justifient l'implantation d'une base sur un "gres système".

2. 2. 2. Tout ce matériel n'est utilisable que s'il existe un "SOFT" correspondant à sa technicité, c'est-à-dire :

- un système opératoire ou OS .
- un moniteur de télétraitement .
- des logiciels spécifiques élaborés .

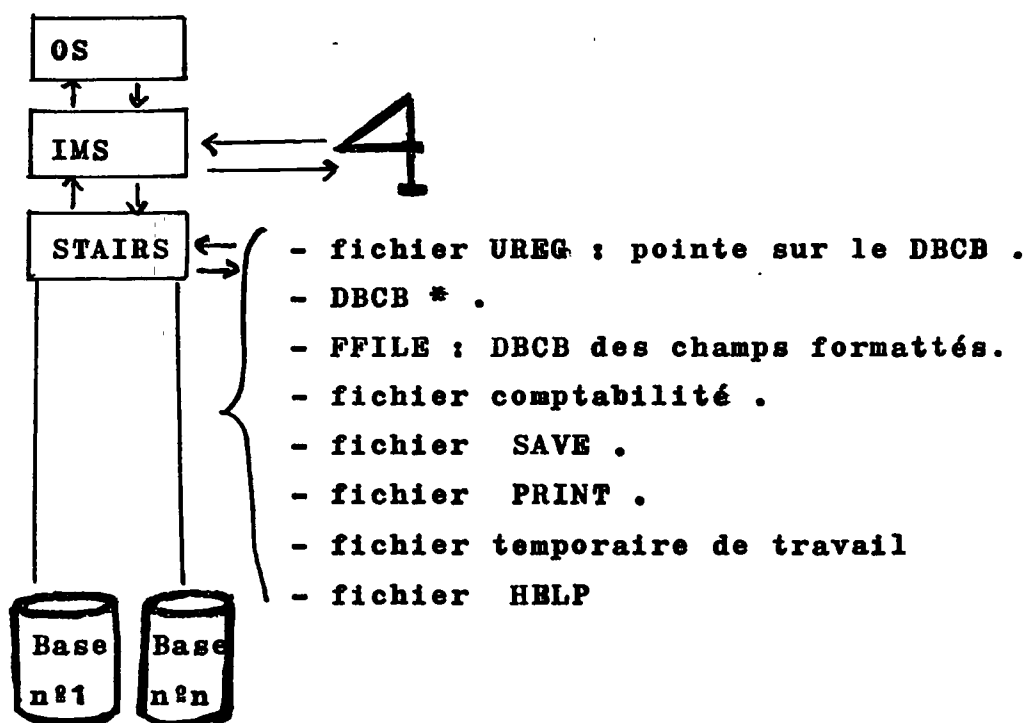
La SPI, pour son activité de documentation en ligne, a choisi le logiciel IBM : STAIRS/IMS (Storage and Information Retrieval System/Information Management System).





2. 2. 3. STAIRS permet de créer et d'emmagasiner des bases de données . Il offre la possibilité de formuler des questions avec différents niveaux de précision; il permet en outre des solutions variées pour l'affichage des documents en sortie . La clause du secret assure l'intégrité du système et des données .

Structure globale de STAIRS à la SPI .



\* Le DBCB (Data Base Control Bloc) contrôle le contenu des paragraphes, le libellé et la description générale de la base .

### **3. IMPLANTATION D'UNE BASE DANS STAIRS .**

### 3. IMPLANTATION D'UNE BASE DANS STAIRS.

#### 3. 1. Définition d'une base dans STAIRS et procédure d'entrée des documents .

##### 3. 1. 1. Une base de données dans STAIRS se compose de quatre fichiers (voir schémas suivants)

- dictionnaire : répertoire des mots .
- inverse : répertoire des occurrences
- indexe : répertoire des champs formatés, protection n° du document .
- texte : documents .

Les documents à entrer dans la base sont mis sous format STAIRS par le programme de reformatage EDIO . Le fichier Pulltext (issu du DBCB) décrit les zones formatées . Le fichier Step indique les mots et les paragraphes à supprimer .

A l'issue de la phase I, on obtient :

- le fichier texte .
- le fichier indexe.
- la liste des mots courts < 15 caractères .
- la liste des mots longs > 15 < 63 caractères .
- les mots courts et longs sont ensuite triés .

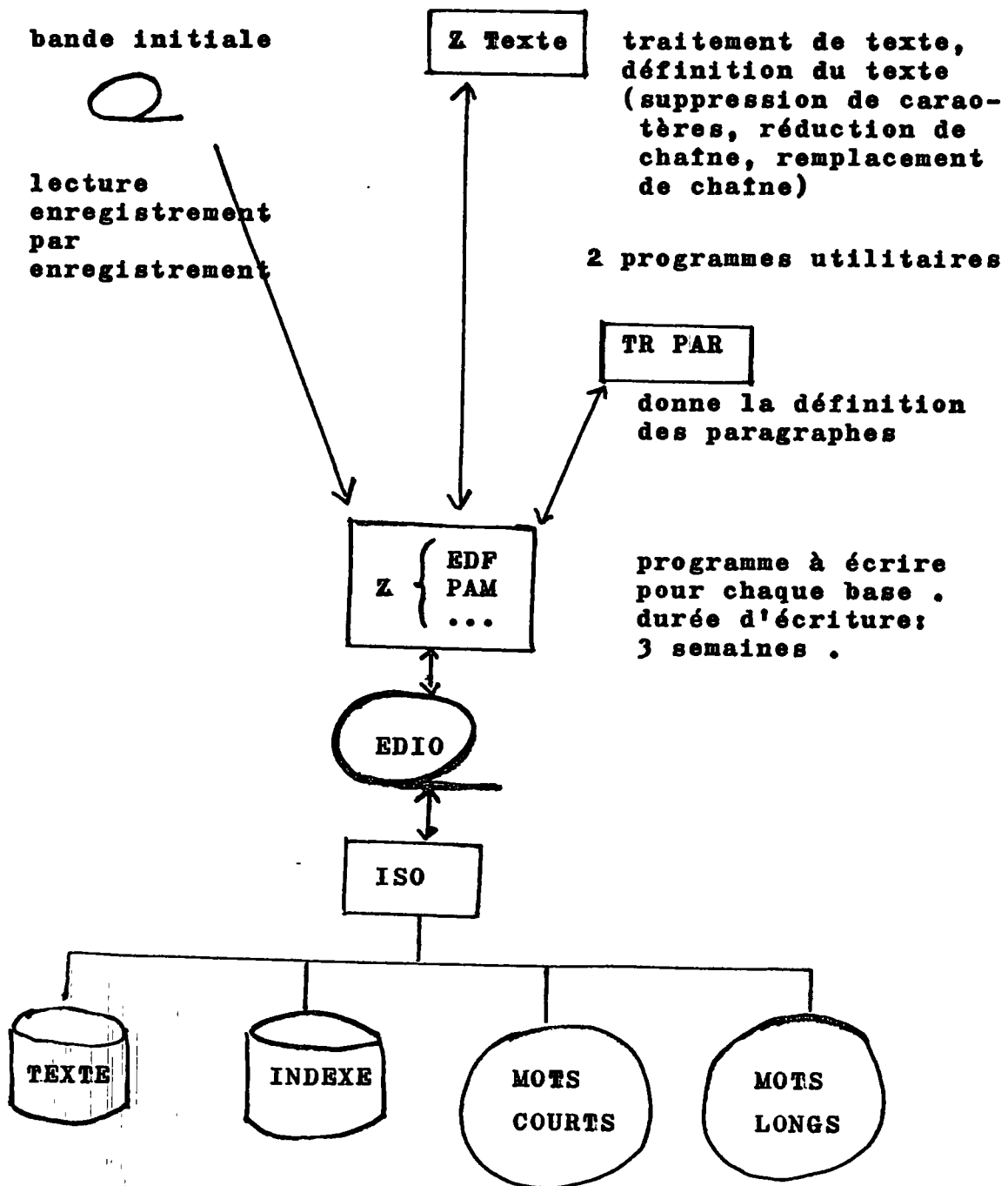
A l'issue de la phase II apparaissent :

- le fichier des mots uniques : les mots sont pris une seule fois . Par exemple, le mot "pollution" qui apparaît n fois ne sera pris qu'une seule fois . Ce fichier donne la liste de tous les mots utilisés .
- le fichier des indicateurs qui donne la position des mots dans le texte .

A l'issue de la phase III sont constitués les fichiers inverse et dictionnaire .

Cependant, le programme de création a été modifié. On obtient la phase I par le biais d'un programme ISO .

En aval, le programme ne change pas, tandis qu'en amont, l'organisation a été modifiée par la SPI .



### 3. 1. 2. Exemple d'un processus de recherche .

Lors d'une recherche en plein texte sous la commande ..SEARCH, l'utilisateur demande un terme qui est recherché dans le fichier dictionnaire . Celui-ci pointe sur le fichier inverse qui fournit le nombre d'occurrences .

Il pointe à son tour sur le fichier indexe qui pointe sur le fichier texte .

fichier dictionnaire :

..SEARCH

matrice

CARTON

accès direct

	A	B	C	D	...
A					
B					
C	+				
D					

le système prend les 2 premières lettres : CA . Il pointe sur une matrice qui pointe à son tour sur CA dans le dictionnaire. Une fois CA trouvé il recherche séquentiellement RTON.

accès direct

AA

:

CA

BLE

LE

RTON

...

accès séquentiel

accès direct

fichier inverse

nombre occurrences	nombre documents	OCC 1	OCC 2	...
		+		

accès direct

+

n° du document    n° de paragraphe, de phrase, de ligne  
n° de mot

fichier indexe

doc. n° 1	champs formatés
doc. n° 2	" "
doc. n° n	" "

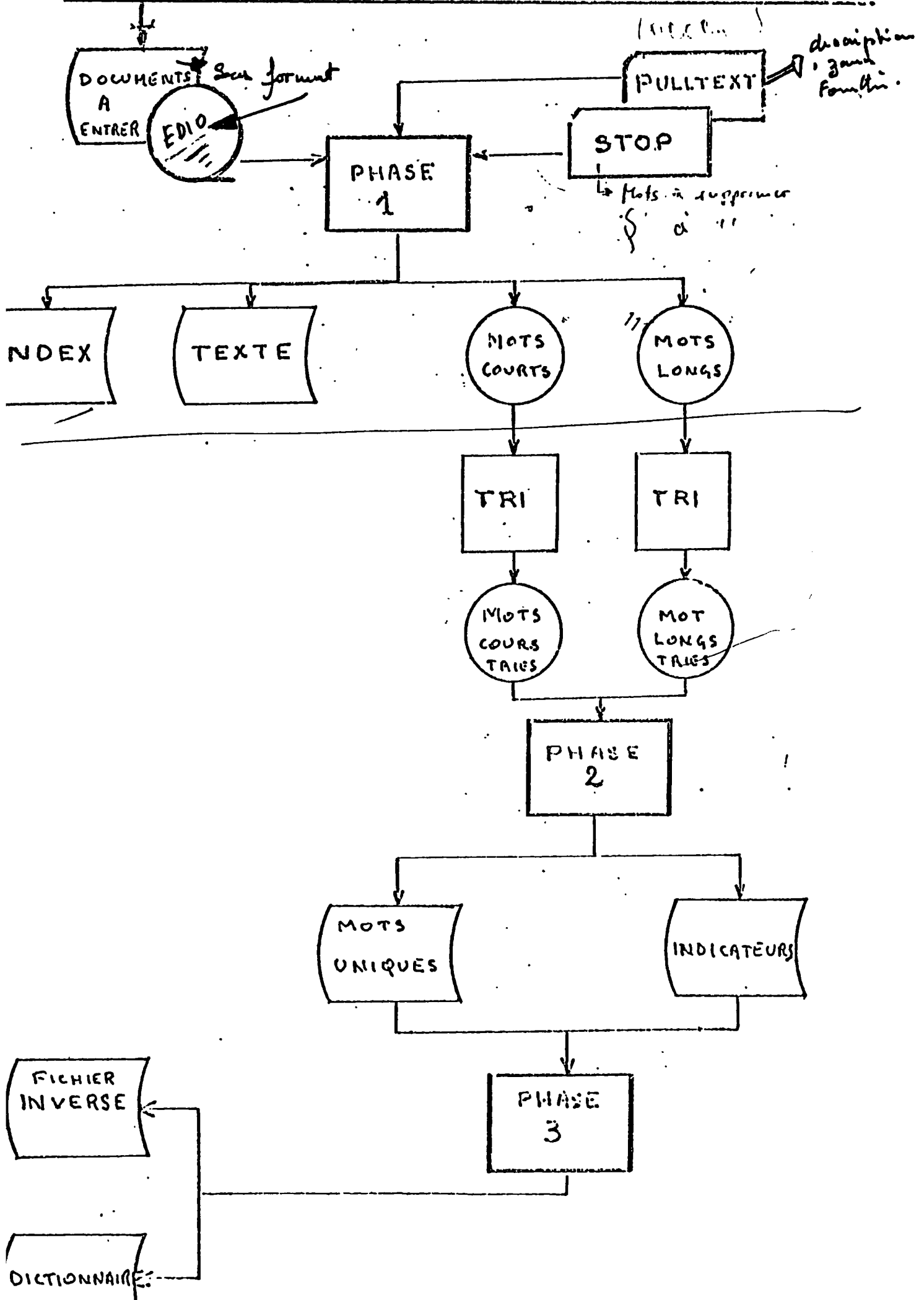
fichier texte

doc. n° 1	texte
doc. n° 2	"
doc. n° n	"

accès direct

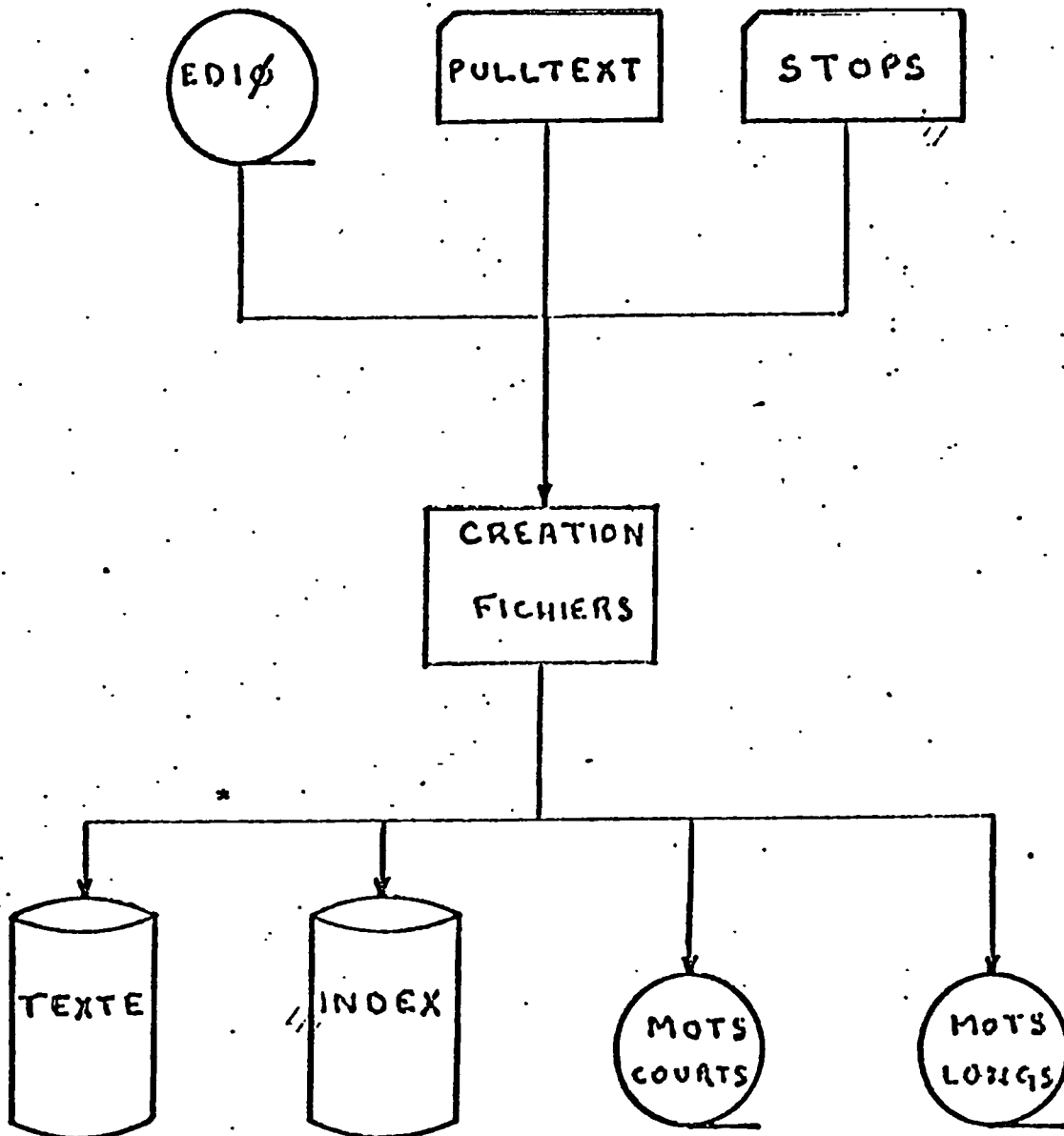


# CREATION D'UNE BANQUE DE DONNEES



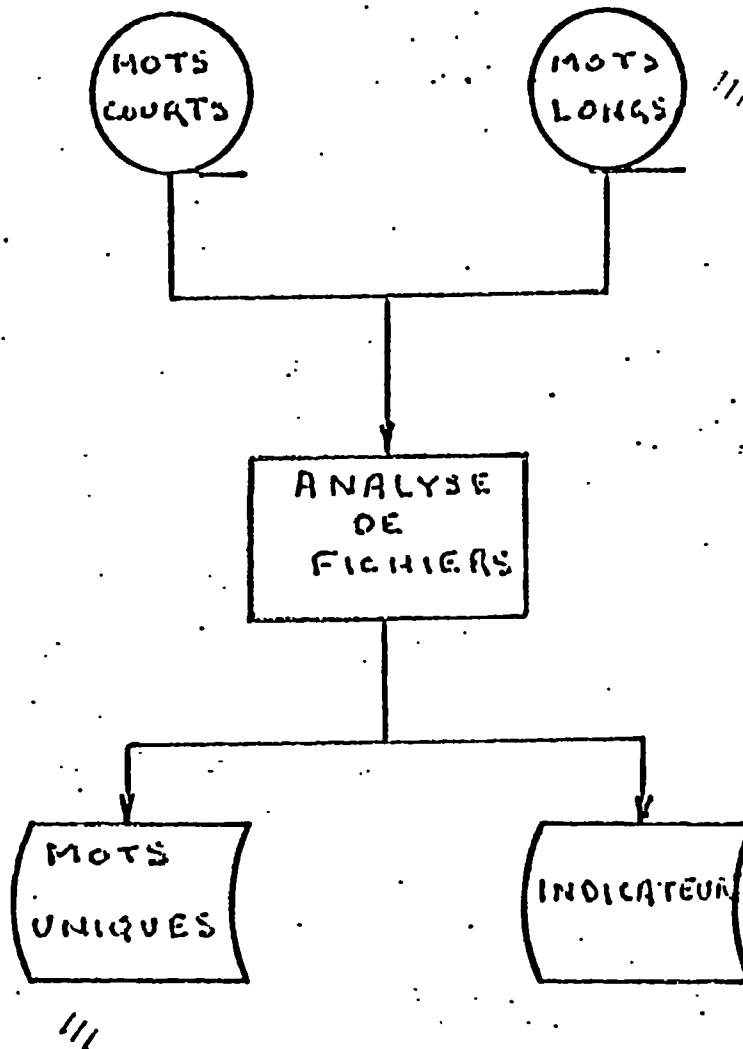
# CREATION D'UNE BANQUE DE DONNEES

## PHASE 1



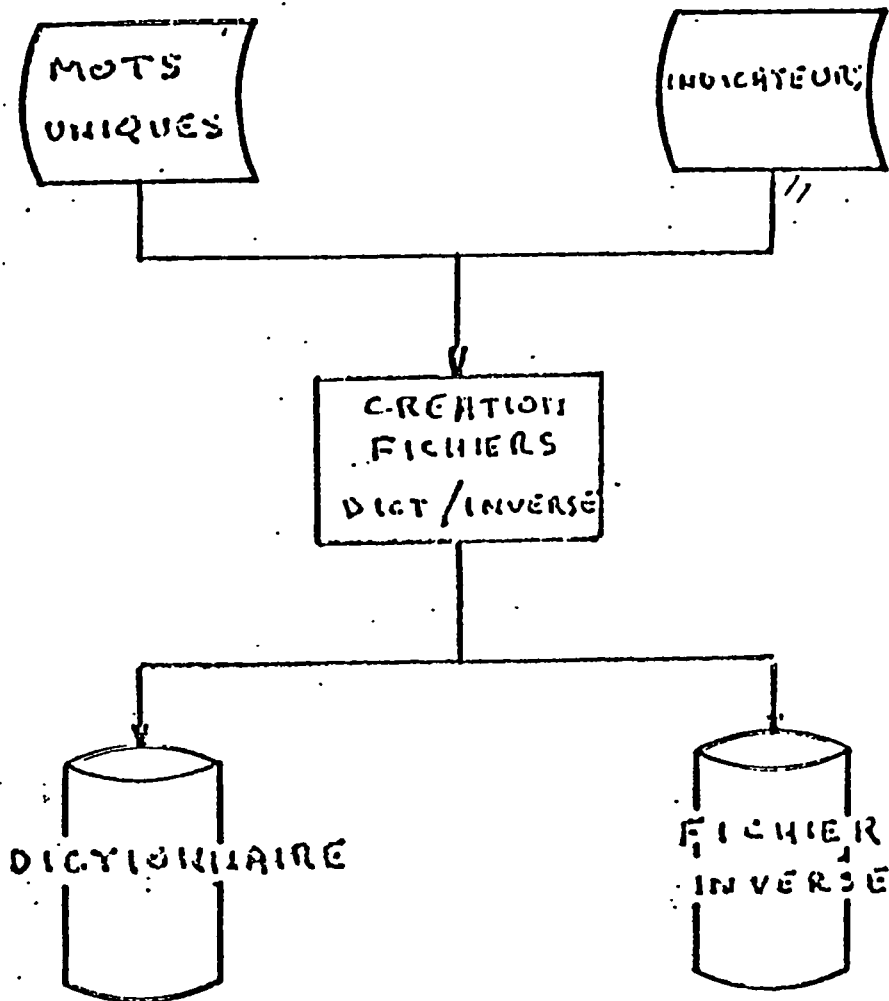
# CREATION D'UNE BANQUE DE DONNEES

## PHASE 2



# CREATION D'UNE BANQUE DE DONNEES

## PHASE 3



## DICTIONNAIRE

EN	LETTER PAIR POINTER							
TETE	LPP 1	LPP 2		LPP 1406				
LETTER PAIR CONTROLS								
LPC 1 ZONES DE CONTROLE	LONGUEURS DE MOTS CUMULEES			MOTS				
	L1	L1+ L2	... L1.... LN	1 2 .... N				
LPC 2 : : LPC 1406								
<table border="1"> <tr> <td>POINTEUR FICHIER INVERSE</td> <td>TOTAUX</td> <td>POINT. SYN.</td> <td>MOT</td> </tr> </table>					POINTEUR FICHIER INVERSE	TOTAUX	POINT. SYN.	MOT
POINTEUR FICHIER INVERSE	TOTAUX	POINT. SYN.	MOT					

## FICHIER INVERSE

ZONE CONTR. E					
Nbr doc.	Nbr occ.	N niveau	PTA %	Occ %	Occ N
.....	occ k				

MAJ min.	Numero relatif DOC.	CODE PAR.	No Phrase	No MOT
-------------	---------------------------	--------------	--------------	-----------

## INDEX

POSITION RELATIVE DOCUMENT	NUMERO DOC. Externe	SUPP.	CODE C N F D	ZONES FORMA- TEES
----------------------------------	---------------------------	-------	--------------------------	-------------------------

## TEXTE

EN TETE DOCUMENT 7	EN TETE PARAG. 1	TEXTE	EN TETE PAR. 2	T E X T E ...
EN TETE PARAG. N	TEXTE	EN TETE DOC. B	EN TETE PARAG. 1	...

### 3. 2. Contraintes STAIRS .

La contrainte la plus importante, quand il s'agit d'implanter une base de données dans STAIRS, est de remettre en forme chaque document en fonction du format STAIRS . Il est nécessaire d'écrire un programme de mise en forme pour chaque base .

Une fois une base implantée, on ne peut pas apporter de modifications . On ne peut corriger un document qu'en le supprimant de la base existante et en rechargeant soit la base complète, soit en utilisant la formule "bases concaténées de STAIRS" (voir paragraphe 3.6.) .

#### Format de présentation d'un document .

##### - Niveau documents :

Le nom du document comporte 12 caractères . Les trois premiers sont alphabétiques et les neuf suivants peuvent être alphanumériques . Un document est constitué d'un ou plusieurs paragraphes .

##### - Niveau paragraphes :

Il existe des classes de paragraphes, qui en regroupent au maximum 16 . Le libellé pour chaque classe de paragraphe doit être inférieur ou égal à 16 caractères .

Le paragraphe est identifié par un code à 3 caractères : le premier est obligatoirement numérique, les autres peuvent être alphabétiques ou alphanumériques . Les paragraphes doivent être en séquence ascendante .

Un paragraphe peut contenir jusqu'à 256 phrases ou 450 lignes de texte .

Une phrase est terminée par les ponctuations : ". " " ? " ou " ! " suivies de 2 blancs .

##### - Niveau phrases :

Les phrases possèdent un numéro qui leur est attribué en fonction de leur apparition dans un paragraphe . Un nouveau paragraphe remet la numérotation à zéro . Une

phrase comporte 256 mots au maximum . Ces numéros de phrases sont utilisés dans l'ordre SEARCH avec l'opérateur SAME .

- niveau mots :

chaque mot est numéroté dans la phrase . Chaque nouvelle phrase remet le compteur à zéro . Dans l'ordre SEARCH ces numéros sont utilisés avec l'opérateur ADJ (pour marquer la notion d'adjacence entre deux mots). Pour le fichier inverse, un mot ne peut pas dépasser 63 caractères, sinon il est tronqué à droite, mais il est conservé dans le texte . Cependant aucun message n'apparaît .

- la ponctuation :

un caractère de ponctuation à la fin d'un mot est supprimé avant que le mot soit placé dans le fichier dictionnaire : (".", ",", ":", ";", "?", "!"). Si le premier ou le dernier caractère d'un mot est un caractère spécial, il est supprimé avant son transfert dans le dictionnaire . Si l'utilisateur veut conserver les caractères spéciaux, il doit placer deux caractères spéciaux au début ou à la fin du mot .

### 3. 3. Les caractères spéciaux .

#### 3. 3. 1. Les caractères spéciaux : aspect technique .

Les caractères spéciaux sont des caractères non alphanumériques restant traductibles .

Exemple : "?" ";" ",." "()" "!" ...

Il existe deux types de caractères :

- les traductibles : lettres et chiffres  
caractères spéciaux
- les non traductibles : code hexadécimal

Sur un total de 256 possibilités, 125 sont utilisées et traduites, il reste donc à peu près 100 caractères intraductibles .

Or la recherche d'un mot dans le fichier dictionnaire se fait par référence à la lettre du début .

Exemple : Dans la recherche du terme Paris, la première lettre étant P, Paris sortira sans problème. Mais on peut avoir :  
 ¥Paris, la recherche se fera alors sur ¥ et non sur P et le mot Paris ne sera pas trouvé .

Pour éviter ces inconvénients, il faut appliquer trois grands principes :

- la transformation d'une chaîne de caractères en une autre (lors du chargement de la base de données).

Exemple : ¥,¥ sera traduit par ,¥

- la suppression ou le remplacement des codes hexadécimaux non traductibles .

Exemple : on remplacera x'04' par ¥

- la réduction d'une chaîne de caractères à un seul élément .

Exemple : ..... ,? on prend le premier caractère rencontré ici "." et on ajoute ¥

Ces trois règles permettent de répondre à la grande majorité des cas de traitement de texte .

### 3. 3. 2. Les caractères spéciaux : comment les normaliser ?

La photocomposition offre une police de caractères très large, tandis que le terminal télétype Texas 745, pris pour base à la SPI, dispose d'un éventail plus succinct de caractères . Il faut prendre une option pour traduire les caractères simples comme le "ç", les accents "^^/¨", les caractères complexes comme "v", "g", "i" ....

Le message est donc obligatoirement appauvri .

Il est important de se mettre d'accord sur les caractères dont disposent les fournisseurs de bases, de se fier à leurs claviers de saisie qui "imposent" les caractères . Il est très difficile de retranscrire les textes photocomposés .

Mais les caractères, qui sortent sur Texas, correspondent-



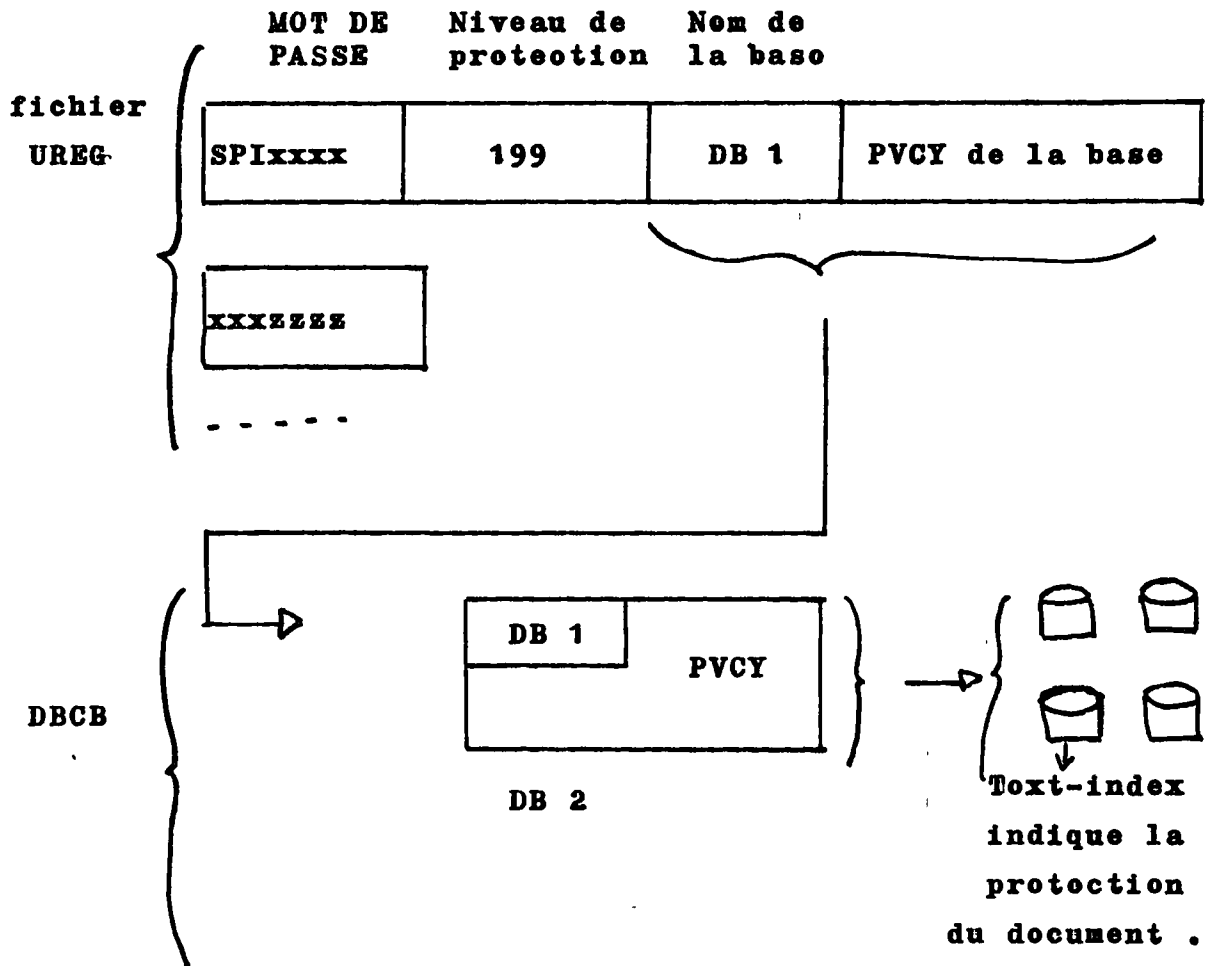
ils à tous les télétypes ? En effet la police du caractère d'un serveur doit être compatible et cohérente avec celle de ses clients . Les touches de service (contrôle S, contrôle H, backspace...) doivent se retrouver sur tous les claviers des clients et être utilisables .

Ce problème est très difficile à résoudre, car il est très vaste, très ardu à cerner .

### 3. 4. Le secret :

3. 4. 1. Une base de données n'est accessible que sous certaines conditions . En effet il faut franchir différents niveaux de protection pour parvenir à l'information .

- Le système d'exploitation demande l'identification du terminal (seulement pour les écrans) .
- AQUARIUS (le module de recherche de STAIRS) demande le mot de passe de l'utilisateur (le "USER"), stocké dans le fichier UREG . Les niveaux d'accès à la base sont aussi définis par le mot de passe dans le fichier UREG. La protection (privacy) générale permet ou non l'effacement de certains documents .



- L'accès à la base est conditionné par le mot de passe de l'utilisateur et par le nom de la base .
- Il existe 256 niveaux de protection différents pour les documents .
- La protection est la même pour tous les paragraphes et les zones formatées d'une base .

### 3. 4. 2. Exemple des protections . (voir pages suivantes)

- exemple 1 : l'utilisateur n'a pas accès au système, car le mot de passe n'a pas été enregistré dans le fichier UREG .
- exemple 2 : l'utilisateur est reconnu; il doit alors ta-

per le nom de la base qu'il veut consulter . Il demande ici à deux reprises une base qui n'est pas accessible par son mot de passe . En revanche, la base A000 est reconnue pour son mot de passe .

La commande ?FILES permet de vérifier les bases en lignes sous le mot de passe .

- exemple 3 : l'utilisateur demande l'affichage du document n°2 . Le "NO CLEARANCE" indique qu'il n'y a pas accès à ce document .

Les documents 5 à 7 ne sont pas visualisables "END OF DOCUMENTS" signale que les 3 documents sont interdits.

- le document DOC = 201 est accessible .

3. 4. 3. Ces différents niveaux de protection assurent un secret absolu tant pour les fournisseurs de bases que pour les utilisateurs .



Exemple II

P READY ♦ SPIDEL READY ♦ SPIDEL READY ♦ SPIDEL READY ♦ SPIDEL READY ♦  
 PIDEL

::06/06/79 ♦♦♦♦ WELCOME ON SPIDEL IT IS : 09.54.11 ♦♦♦♦

0102 ENTER DATA BASE NAME\_:

120

0103 USER NOT REGISTERED FOR DATA BASE, RE-ENTER DATA BASE NAME\_:

130

0103 USER NOT REGISTERED FOR DATA BASE, RE-ENTER DATA BASE NAME\_:

000

AQUARIUS - SEARCH MODE - BEGIN YOUR QUERY AFTER THE STATEMENT NUMBER

0001\_:

FILES

000 DATA BASE CETIM

030 DATA BASE CIS

010 DATA BASE EDF

.OFF

\*\*\*\*\* COMPTABILITE DE VOTRE SESSION \*\*\*\*\*

BASE=A000 HEURE-DEBUT:09.54.12 HEURE-FIN:09.55.38

DUREE DE VOTRE SESSION: 1MN.

\*\*\*\*\*

0707 ♦ STAIRS/VS - AQUARIUS SESSION COMPLETE.

Exemple III

:  
 .S  
 QUARIUS - SEARCH MODE - BEGIN YOUR QUERY AFTER THE STATEMENT NUMBER  
 0001:  
 .PAGE LINES=99  
 :  
 ABLE\$1 ADJ COAXIS

RESULT 324 OCCURRENCE 201 DOCUMENTS

:00002:  
 .B 1 ALL -> demandes d'affichage  
 0521 \* ENTER DOCUMENT REQUEST OR ANOTHER COMMAND

DC=2  
 0502 \* NO CLEARANCE FOR DOCUMENT.  
 0521 \* ENTER DOCUMENT REQUEST OR ANOTHER COMMAND

DC=5,7

END OF DOCUMENTS

DC=201  
 DF79H512302 DOCUMENT= 201 OF 201 NUMBER OF LINES = 30  
 NUM = 79H512302  
 DATE = 197711  
 LANG = FR  
 CODEM = SEE

ITD ETUDE DES CABLES COAXIAUX A TRESSE MULTIPLE SOUMIS A DES  
 PERTURBATIONS ELECTROMAGNETIQUES  
 JT DEGAUQUE P. DEMOULIN B.  
 JU SOURCE : CONGRES SEE, GIF - SUR - YVETTE, 16 / 11 / 1977, 7 PAGES,  
 5 FIGURES, 8 REF. BIBL. 1977 / 11 / 16  
 EN EN FRANCAIS, CODE SOURCE : SEE, REFERENCE : COMMUNICATION 7725422,  
 ACCESSIBILITE : LIBRE, TYPE D'INFORMATION : COMMUNICATION DE  
 CONGRES, COMMENTAIRES : METHODES NUMERIQUES EN PROTECTION  
 ELECTROMAGNETIQUES  
 ES ANALYSE THEORIQUE, UTILISANT LA NOTION D'IMPEDANCE DE TRANSFERT,  
 DE L'EFFICACITE DU BLINDAGE DE CABLES COAXIAUX A TRESSES MULTIPLES,  
 SOUMIS A DES TENSIONS PERTURBATRICES ; CALCUL NUMERIQUE ( PAR  
 ORDINATEUR ) DE CETTE EFFICACITE, EN FONCTION DES PARAMETRES  
 PHYSIQUES DES BLINDAGES EMPLOYES. PRESENTATION DE LA TECHNIQUE D'  
 ANALYSE EMPLOYEE, FORMULATION DES ALGORITHMES. ETUDE PARAMETRIQUE,  
 SUR CES BASES, DU CABLE A TROIS TRESSES. DISCUSSION DE L'  
 APPLICABILITE DE CETTE APPROCHE AU CALCUL DES PERFORMANCES DES  
 BLINDAGES.  
 LE CABLE ELECTRIQUE. LIGNE DE TRANSMISSION. CABLE COAXIAL. CABLE  
 BLINDE. TRESSE. CALCUL. PROTECTION. PROTECTION ELECTRIQUE.  
 THEORIE. MODELISATION. ANALYSE FONCTIONNELLE. METHODE.  
 RESOLUTION NUMERIQUE. PERFORMANCE. FRANCE. 1977. IMPEDANCE DE  
 TRANSFERT. PERTURBATION.

0601 \* END OF DOCUMENTS IN LIST - ENTER RETURN OR ANOTHER COMMAND.

:  
 .OFF  
 \*\*\*\*\* COMPTABILITE DE VOTRE SESSION \*\*\*\*\*  
 BASE=A010 HEURE-DEBUT:10.00.22 HEURE-FIN:10.06.15  
 DUREE DE VOTRE SESSION: 6MN.  
 \*\*\*\*\*  
 0707 \* STAIRS/VS - AQUARIUS SESSION COMPLETE.

### 3. 5. Chargement et contrôle d'une base .

On procède en premier lieu au chargement d'environ 500 documents pour voir si les résultats obtenus correspondent aux demandes des fournisseurs . (annexe 3.5.) .

Une série de tests est alors effectué pour relever les erreurs possibles :

- un affichage des zones formatées afin de se rendre compte si la longueur désirée est respectée, si la zone est bien numérique ou alphanumérique .
- un affichage des paragraphes des zones libres pour voir si les libellés sont conformes à ceux du fournisseur et s'ils correspondent aussi au schéma STAIRS, afin que les bases se présentent d'une façon homogène .
- une recherche en plein texte sur une large question . L'utilisation des opérateurs booléens .
- un "Select" est ensuite fait sur toutes les zones formatées pour tester leur validité .
- un test sur les caractères spéciaux .  
Par exemple un problème était apparu avec la base CETIM qui utilisait comme séparateur dans les mots clés le "%" . La SPI a fait alors disparaître ces séparateurs; ou il existait aussi des % dans les autres paragraphes qui furent aussi détruits à cette occasion .
- regarder s'il existe des mots vides, qui allègent la base . Mais le problème de la suppression systématique des articles par exemple est très délicat . En effet tous les articles : le, la, les, l' étaient pris comme mots vides pour le CIS; or le CIS se servait des lettres A...L de l'alphabet pour désigner la nature des documents . Donc les problèmes de mots vides varient selon les bases .
- la confidentialité est aussi fortement testée .

Une base comme l'EDF comporte 10% de documents confidentiels et une erreur de la part du serveur peut être très préjudiciable .

- toutes les fonctions sont testées .

- une demande d'impression d'environ 200 documents est commandée pour pouvoir examiner sur un large éventail les défauts qui peuvent apparaître .

La série de tests complète dure 1 journée 1/2

Une fois les tests terminés, la base chargée est donnée au fournisseur pour qu'il l'examine à son tour . S'il en est satisfait, il permet le chargement du reste de la base . On vérifie ensuite par une même série de tests si un chargement plus abondant a pu susciter d'autres problèmes, si le nombre de documents enregistrés est conforme à ce qui est prévu.

Cette opération demande une 1/2 journée de tests.

La base est de nouveau mise en test chez le fournisseur qui, s'il n'a pas de modification à faire, donne la permission de commercialiser .

Le chargement est un problème délicat qui demande des délais très variables .

### 3. 6. Les mises à jour . La sauvegarde des fichiers .

#### 3. 6. 1. Les mises à jour .

Elles se font en batch, d'une façon périodique, environ tous les mois . Elles dépendent de la volonté des fournisseurs de bases .

Elles demandent des manipulations importantes et des précautions sortant de l'ordinaire .

Une base se constitue de 4 fichiers : index,

inverse,

texte,

dictionnaire .



### Principes de la mise à jour .

À partir d'une base est stockée dans un fichier appelé : suffixe . À chaque mise à jour on rajoute un suffixe . À la quatrième fois, on fusionne les 4 suffixes et on redémarre à 1 .

Solution retenue pour la conduite du traitement : le lancement est délicat parce que les programmes de fusion doivent connaître les caractéristiques techniques de la base du nouveau fichier à créer, pour réserver un espace disque suffisant pour effectuer le travail . Or d'une base à l'autre la taille du document varie .

La solution retenue a pour but de supprimer la détermination manuelle des paramètres . Pour cela un programme (en Cobol) génère les cartes J.C.L. de lancement avec leurs paramètres calculés automatiquement (numéro de suffixe, place disque à réserver).

### 3.6.2. Les problèmes de sécurité .

On crée une nouvelle base, sans rien toucher au reste . On ne touche, ni ne détruit les anciens suffixes . Dès que le nouvel exemplaire est bon, il est mis en ligne . On ne fait ainsi aucune sauvegarde particulière avant la mise à jour . Si le résultat est mauvais, on recommence l'opération .

Le DBCB (Data Base Control Bloc) effectue le dernier travail : l'obtention d'un exemplaire concret : le suffixe . Quand il y a fusion en un nouveau suffixe sur bande, on déverse le contenu de la bande sur le suffixe 1 et on vide les autres . La sécurité est automatique . En effet la procédure de fusion se fait sur bande et non sur le fichier disque . On conserve donc toujours l'ancienne version, et en cas d'échec de la procédure, les anciens fichiers sont repris . Ainsi, il n'est pas nécessaire de sauvegarder les fichiers à fusionner .

### 3. 6. 3. La sécurité sur l'ensemble des bases .

Il existe une sauvegarde générale sur disque :

- une fois par semaine, les disques contenant les bases sont sauvegardés .

Si des difficultés surviennent :

- avant une mise à jour : on reprend la sauvegarde ,
- après une mise à jour : on doit refaire la dernière mise à jour à partir de la sauvegarde et de la bande de mise à jour .

Les données des mises à jour :

- sur une bande, on garde la dernière mise à jour jusqu'à ce que la suivante ait été effectuée et réussie .
- en archive en double, sur bande, toutes les données d'une base, depuis le début, ce qui peut être utile par la suite pour un éventuel reformattage complet de la base .
- on demande au fournisseur d'être en mesure de refournir sa bande de mise à jour .

### 3. 6. 4. Simulation de la mise à jour d'une base de données (documents I et II).

Le document I représente :

- dans la situation ancienne : l'existence de 3 suffixes pleins ,
  - dans la situation nouvelle : le chargement du 4ème suffixe avec les 4000 documents de la dernière mise à jour .
- Il est nécessaire de connaître le nombre de documents pour définir l'espace disque à utiliser .

Or, quand les 4 suffixes sont occupés, il faut faire la fusion des 4 suffixes pour n'occuper que le premier : Document II.

Ainsi, dans la situation ancienne,  
l'addition du suffixe 1 = 1000 documents

"	"	2 = 2000	"
"	"	3 = 3000	"
"	"	4 = 4000	"

donne 10000 documents, qui se retrouvent dans le suffixe 1 de la situation nouvelle . Les trois autres suffixes sont libérés pour de nouvelles mises à jour .

## SITUATION ANCIENNE

//	BASE=00,	NUMERO DE LA BASE	00000010
//	LR=00300,	LONGUEUR DU RECORD TEXTE INDEX	00000020
//	BL=01200,	LONGUEUR DU BLOC TEXTE INDEX	00000030
//	SPTNDX=000000,	PLACE TEXTE INDEX= NB DOC * 010 / 0002	00000040
//	SPTXT=000000	PLACE TEXTE = NB DOC * 005 / 0010	00000050
//	(SUF=3	SUFFIXE BASE = 1 2 3 OU 4	00000060
//	SPINVT=000000	PLACE INVERSE = NB DOC * 032 / 0004	00000070
//	SPDICT=000000	PLACE DICT = NB DOC * 004 / 0020	00000080
//	NBDNC=000000	NOMBRE DE DOCUMENTS	00000090
//	DATE=000000	DATE DU CHARGEMENT	00000100
//	NBSUF1=001000	NB DOC SUF1	00000110
//	NBSUF2=002000	NB DOC SUF2	00000120
//	NBSUF3=003000	NB DOC SUF3	00000130
//	NBSUF4=000000	NB DOC SUF4	00000140

## SITUATION NOUVELLE

//	BASE=00,	NUMERO DE LA BASE	00000010
//	LR=00300,	LONGUEUR DU RECORD TEXTE INDEX	00000020
//	BL=01200,	LONGUEUR DU BLOC TEXTE INDEX	00000030
//	SPTNDX=020001,	PLACE TEXTE INDEX= NB DOC * 010 / 0002	00000040
//	SPTXT=002001	PLACE TEXTE = NB DOC * 005 / 0010	00000050
//	(SUF=4	SUFFIXE BASE = 1 2 3 OU 4	00000060
//	SPINVT=032001	PLACE INVERSE = NB DOC * 032 / 0004	00000070
//	SPDICT=000801	PLACE DICT = NB DOC * 004 / 0020	00000080
//	NBDNC=004000	NOMBRE DE DOCUMENTS	00000090
//	DATE=11/05/79	DATE DU CHARGEMENT	00000100
//	NBSUF1=001000	NB DOC SUF1	00000110
//	NBSUF2=002000	NB DOC SUF2	00000120
//	NBSUF3=003000	NB DOC SUF3	00000130
//	NBSUF4=004000	NB DOC SUF4	00000140

DOO. 1

## SITUATION ANCIENNE

//	BASE=00,	NUMERO DE LA BASE	00000010
//	LR=00300,	LONGUEUR DU RECORD TEXTE INDEX	00000020
//	RL=01200,	LONGUEUR DU BLOC TEXTE INDEX	00000030
//	SPTNDX=000000,	PLACE TEXTE INDEX= NB DOC * 010 / 0002	00000040
//	SPTXT=000000	PLACE TEXTE = NB DOC * 005 / 0010	00000050
//	SUF=④	SUFFIXE BASE = 1 2 3 CU 4	00000060
//	SPINVT=000000	PLACE INVERSE = NB DOC * 032 / 0004	00000070
//	SPDICT=000000	PLACE DICT = NB DOC * 004 / 0020	00000080
//	NBDOC=000000	NOMBRE DE DOCUMENTS	00000090
//	DATE=000000	DATE DU CHARGEMENT	00000100
//	NBSUF1=001000	NB DOC SUF1	00000110
//	NBSUF2=002000	NB DOC SUF2	00000120
//	NBSUF3=003000	NB DOC SUF3	00000130
//	NBSUF4=004000	NB DOC SUF4	00000140

## SITUATION NOUVELLE

//	BASE=00,	NUMERO DE LA BASE	00000010
//	LR=00300,	LONGUEUR DU RECORD TEXTE INDEX	00000020
//	BL=01200,	LONGUEUR DU BLOC TEXTE INDEX	00000030
//	SPTNDX=050001,	PLACE TEXTE INDEX= NB DOC * 010 / 0002	00000040
//	SPTFXT=005001	PLACE TEXTE = NB DOC * 005 / 0010	00000050
//	SUF=①	SUFFIXE BASE = 1 2 3 CU 4	00000060
//	SPINVT=080001	PLACE INVERSE = NB DOC * 032 / 0004	00000070
//	SPDICT=002001	PLACE DICT = NB DOC * 004 / 0020	00000080
//	NBDOC=010000	NOMBRE DE DOCUMENTS	00000090
//	DATE=11/05/79	DATE DU CHARGEMENT	00000100
//	NBSUF1=010000	NB DOC SUF1	00000110
//	NBSUF2=000000	NB DOC SUF2	00000120
//	NBSUF3=000000	NB DOC SUF3	00000130
//	NBSUF4=000000	NB DOC SUF4	00000140

DOC. II

#### 4. GESTION D'UN SERVEUR .

#### 4. GESTION D'UN SERVEUR .

Le service SPIDEL doit mettre en place un système de gestion d'un service "grand public" . En effet, assurer un rôle de serveur implique, non seulement le choix judicieux des bases, une infrastructure solide et une bonne équipe de techniciens, mais aussi la maîtrise de certains problèmes, en particulier :

- à qui, quand et comment envoyer de la documentation et quel type de documentation, d'où la nécessité d'un fichier adresse .
- gérer la documentation des fournisseurs de bases de données, et la documentation payante SPIDEL .
- la facturation .

##### 4. 1. Le fichier adresse .

(voir annexe 4.1.)

Deux personnes s'occupent de la promotion de SPIDEL en France et à l'étranger .

Des centaines de personnes, choisies selon leur secteur d'activité, reçoivent un courrier de présentation SPIDEL. Si les destinataires en manifestent le désir, une documentation complète leur est envoyée . Ils peuvent même assister à une journée de présentation .

Le suivi des clients doit se faire régulièrement . Il faut donc être en mesure de discerner les différentes catégories de personnes et leur envoyer la documentation correspondant à leurs besoins . Ceci suppose donc l'aide de l'informatique .

Le fichier adresse demande une attention particulière :

- éliminer les doublons .
- corriger les adresses et les noms inexacts.
- normaliser les sigles .

Le tri s'effectue actuellement sur :

- le code postal
- le responsable commercial
- le nom de la société

- le nom de la personne

Toutes les corrections se font en temps réel sous le TSO  
(Logiciel IBM)

Des étiquettes d'adresses sortent après traitement en  
batch .

#### 4. 2. La documentation payante .

A l'heure actuelle, les commerciaux s'occupent manuellement :

- de la gestion de la documentation des fournisseurs :  
les thesauri .
- de la vente des manuels d'utilisation STAIRS .
- de la vente des manuels de contenu des bases .

#### 4. 3. Facturation .

4. 3. 1. Les problèmes de facturation sont assez ardues,  
à cause de régimes particuliers appliqués à certains  
clients .

Il faut en effet tenir compte :

- des tarifs normaux .
- des tarifs particuliers (par exemple les abonnés au CIS)
- les bénéficiaires d'heures gratuites .
- des reports de facturation : si une somme n'est pas  
assez élevée pour un mois, le client n'est facturé qu'au  
bout du trimestre .

#### 4. 3. 2. Les réseaux de communication .

TRANSPAC va facturer aux serveurs le temps de connection  
des utilisateurs . Le serveur doit à son tour répartir les  
dépenses .

La gestion du service SPIDEL, tout en tenant compte de ces  
nombreux paramètres, ne doit pas empiéter trop largement  
sur les bénéfices du service .

## C O N C L U S I O N

L'activité SPIDEL, avec ses trois bases, le CETIM, l'EDF et le CIS, répond donc à des besoins réels d'information dans le domaine économique et technique . Le logiciel STAIRS permet d'autre part une interrogation facile de ces bases .

L'implantation d'une base sur un serveur demande cependant beaucoup d'attention et de soins . En effet, décider de choisir telle base plutôt que telle autre implique un risque si la base ne répond pas à l'attente de l'utilisateur .

Si SPIDEL dispose de moyens informatiques très suffisants, il doit néanmoins veiller à présenter aux utilisateurs un service d'une qualité parfaite . Il doit sans cesse contrôler les bases, s'assurer de leur mise à jour, de la sécurité des documents et du secret de l'information. La gestion, tout en demandant de la souplesse, doit être menée avec beaucoup de fermeté .

Ainsi, un serveur doit toujours aller de l'avant et être prêt à se remettre en question en acceptant les remarques de l'extérieur . Il doit aussi être sans cesse sur la défensive et s'adapter à l'environnement économique et politique d'un pays .



## ANNEXES

## SOMMAIRE

1. FOURNISSEUR
  - 1.1 ADRESSE - TÉLÉPHONE - TÉLEX
  - 1.2 INFORMATION SUCCINCTE
  - 1.3 PERSONNE A CONTACTER
2. CONTENU
  - 2.1 TYPE DE DOCUMENTS
  - 2.2 ORIGINE DES INFORMATIONS
  - 2.3 NOMBRE DE DOCUMENTS DU FICHIER
  - 2.4 PÉRIODE DE TEMPS COUVERTE
  - 2.5 PÉRIODICITÉ DES MISES A JOUR
3. DOMAINES COUVERTS
  1. PARAGRAPHES ET RUBRIQUES D'UN DOCUMENT
    - 1.1 DESCRIPTION DÉTAILLÉE DE CHAQUE RUBRIQUE ET CHAQUE PARAGRAPHE
    - 1.2 SIGNALEMENT D'UN DOCUMENT
4. METHODE D'ACCES
  - 4.1 THÉSAURUS - UTILISATION
  - 4.2 MOTS-VIDE ; SYNONYMES
  - 4.3 ABRÉVIATIONS; PONCTUATION PARTICULIÈRE

- 5.4            CARACTÈRES SPÉCIAUX
- 5.5            DOCUMENTS SPÉCIAUX
- 6.            DOCUMENTS ORIGINAUX
- 6.1            CONDITIONS DE SECRET
- 7.            OUTILS POUR L'UTILISATION DE LA BASE
- 8.            EXTRAIT DU THESAURUS
- 9.            LISTE DES DOCUMENTS ANALYSES

SOCIETE POUR L'INFORMATIQUE  
SERVICE CONDUITE  
98, Boulevard Victor Hugo  
92115 - CLICHY  
A l'attention de M. IMBERT

## FEUILLE DE LIAISON

Veillez trouver, ci-jointe, une bande magnétique de mise à jour dont les références (à inscrire également sur le support) sont :

CODE BASE ..... : [ ] [ ]  
DATE D'ENVOI ..... : [ ] [ ] / [ ] [ ] / [ ] [ ]  
NOMBRE DE DOCUMENTS ..... : [ ] [ ] [ ] [ ] [ ] [ ]

Signature : ..

NOTA :

1 - Les caractéristiques techniques suivantes ne doivent pas être changées sans l'accord de la SPI :

2 - La bande magnétique doit voyager dans une enveloppe spéciale munie d'une étiquette d'expédition écrite recto (adresser aller) et verso (adresse retour) fournie par la SPI.

La bande magnétique est à renvoyer sous 48 heures accompagnée de la feuille ci-jointe

## FEUILLE DE RETOUR

SOCIETE POUR L'INFORMATIQUE  
SERVICE CONDUITE

Veillez trouver en retour la  
bande magnétique relative à votre envoi  
du [ ] [ ] / [ ] [ ] / [ ] [ ]

Date : \_\_\_\_ / \_\_\_\_ / \_\_\_\_ Signature : ..

IDENTIFICATION DE LA BASE

Code HTA Libelle'

FOURNISSEUR DES MISES A JOUR

Adresse

Téléphone Telex

Personne a' contacter

PERIODICITE DES MISES A JOUR

Fréquence Date approximative d'envoi des NaJ

DESCRIPTION TECHNIQUE DE LA BANDE DE MISE A JOUR

( TI/SE doit s'efforcer d'obtenir du fournisseur une Bande en standard label IBM)

DSN Densité'

RECFM LRECL BLKSIZE

LABEL

( Ces caractéristiques ne pourront pas être modifiées sans l'accord de RE/cp )

OPERATIONS FAITES PAR TI/SE SUR LA BASE AVANT PRISE EN CHARGE PAR RE/CP

Date	Opération*	Suffixe créé	Nombre de documents	Observations
/				
/				
/				
/				
/				
/				

\* Changement initial, Mise à jour ou fusion

1	//		16		31		59		65		73
	//	BASE=   ,				NUMERØ DE LA BASE	-----				00000010
	//	LR=   ,				LØNGUEUR DU RECØRD TEXTE INDEX	(1,3)				00000020
	//	BL=   ,				LØNGUEUR DU BLØC TEXTE INDEX					00000030
	//	SPTNDX=000000,				PLACE TEXTE INDEX = NB DØC *     /     +1	(4,3)	(4,2)			00000040
	//	SPTEXT=000000,				PLACE TEXTE = NB DØC *     /     +1	(5,3)	(5,2)			00000050
	//	SUF=   ,				SUFFIXE BASE = 1 2 3 OU 4					00000060
	//	SPINVT=000000,				PLACE INVERSE = NB DØC *     /     +1	(8,3)	(8,2)			00000070
	//	SPDICT=000000,				PLACE DICT = NB DØC *     /     +1	(8,1)	(8,2)			00000080
	//	NBDØC=000000,				NØMBRE DE DØCUMENTS DU SUF EN CHARGE					00000090
	//	DATE=00/00/00				DATE DU CHARGEMENT					00000100
	//	NBDØC 1=   ,				NØMBRE DE DØCUMENTS DU SUFFIXE 1					00000110
	//	NBDØC 2=   ,				NØMBRE DE DØCUMENTS DU SUFFIXE 2					00000120
	//	NBDØC 3=   ,				NØMBRE DE DØCUMENTS DU SUFFIXE 3					00000130
	//	NBDØC 4=   ,				NØMBRE DE DØCUMENTS DU SUFFIXE 4					00000140

Remplir sans lasser de l'aine toutes les zones numérotées qui doivent toutes être numériques.

SP1 47

membrane DBAL dans la HT. ST2. CONTLIB

DELA  
ADDA  
A

DSENDI

10000001

# I - STRUCTURE PHYSIQUE DE LA BANDE

\* LA BANDE EST-ELLE IBM

OUI - NON

\* DENSITE (800, 1 600, 6 250 bpi)

\* IL Y-A-T'IL UN LABEL ( au sens IBM)

OUI - NON

\* NOM DE LA BANDE OU DSN

\* LABEL STANDARD

OUI - NON

\* LABEL UTILISATEUR

OUI - NON

\* VOL - SER

\* FORMAT ENREGISTREMENT  
(ou RECFM)

FIXE - FIXE BLOC  
VARIABLE - VARIABLE BLOC  
INDEFINI

\* LONGUEUR ENREGISTREMENT  
(ou LRECL)

\* INDIQUER LONGUEUR MAXIMUM  
EN CAS LONGUEUR VARIABLE OU INDEFINIE

\* LONGUEUR BLOC SI FORMAT BLOQUE  
(ou BLKSIZE)

\* INDIQUER LONGUEUR MAXIMUM  
EN CAS DE LONGUEUR VARIABLE OU INDEFINIE



## II - STRUCTURE LOGIQUE DE LA BANDE

### 1) INFORMATIONS STATISTIQUES

* DOCUMENT	{	NBRE DE DOCUMENTS	:	<input type="text"/>
	{	TAILLE MAXIMUM D'1 DOCUMENT	:	<input type="text"/>
* PARAGRAPHE	{	NBRE MAXIMUM DE PARAGRAPHES	:	<input type="text"/>
	{	TAILLE MAXIMUM D'UN PARAGRAPHE	:	<input type="text"/>
* MISE À JOUR	{	NBRE MAXIMUM DE DOCUMENTS	:	<input type="text"/>
	{	PÉRIODICITÉ DE LA MISE À JOUR	:	<input type="text"/>
* COMMENTAIRES	:			

### 2) TRI

* LA BANDE EST-ELLE TRIÉE ?	:	OUI - NON
* SUR LE NUMÉRO DE DOCUMENT ?	:	OUI - NON
* DANS UN DOCUMENT SUR LE N° DE PARAGRAPHE ?	:	OUI - NON
* CRITÈRE DE TRI (INDIQUER LA POSITION ET DÉPLACEMENT DANS L'ENREG.)	:	<input type="text"/>
* COMMENTAIRES :		

- 5 -

### 3) DESCRIPTION DES ENREGISTREMENTS

\* POUVEZ-VOUS FOURNIR EN CLAIR UN EXEMPLE D'ÉDITION DES  
ENREGISTREMENTS DE LA BANDE ( DUMP ) ?

\* POUVEZ-VOUS FOURNIR UN F.L.E ?  
(FORMAT LOGIQUE D'ENTRÉE)

\* DESCRIPTION DÉTAILLÉE DES ENREGISTREMENTS :

\* COMMENTAIRES :

### III - TRAITEMENT DE LA BANDE

#### 1) DOCUMENT

\* LE N° DE DOCUMENT DOIT-IL ÊTRE GÉNÉRÉ PAR NOTRE SOCIÉTÉ ( S.P.I ) ?

: OUI - NON

PRÉCISER OU LE TROUVER :

\* SI NON, DANS QUEL PARAGRAPHE ?

SUR QUELLE LONGUEUR ?

A QUELLE POSITION ?

\* FAUT-IL TRAITER TOUS LES DOCUMENTS ?

: OUI - NON

\* SI NON, PRÉCISER LE CRITÈRE DE SELECTION :

\* PROTECTION AU NIVEAU DU DOCUMENT ?

\* COMMENTAIRES :

(1) faire une croix si on désire la visualisation

[illegible]

3) TRAITEMENT DES CARACTÈRES (caractères spéciaux, ponctuation, transcodage)

- CARACTÈRES SPÉCIAUX

\* LES CARACTÈRES SONT-ILS EN CODE EBCDIC ? OUI - NON

\* IL Y-A-T-IL DES CARACTÈRES NON ÉDITABLES ? OUI - NON

- PONCTUATION

\* QUELS SONT LES CARACTÈRES DE PONCTUATION ?

- RÈGLES DE PONCTUATION (mettre une croix dans la colonne remplissant cette condition)

CARACTERE	Un espace avant ce caractère (1)	Un espace après ce caractère (1)	Espace avant ou après (1)	SUPPRESSION (1)	TRANSCODAGE	COMMENTA

COMMENTAIRES :

IV - RENSEIGNEMENTS COMPLÉMENTAIRES

N° TÉLÉPHONE

NOMS	SOCIÉTÉ	TÉL. DE LA SOCIÉTÉ	POSTE
Mr MEREL	S.P.I.	739-94-40	508

2) COMMENTAIRES :

## ANNEXE 4.1.

VOS REF.  
NOS REF. LT120JCS

CLICHY, le 7 Juin 1979

Madame, Monsieur,

Nous avons le plaisir de vous faire part du lancement de SPIDEL, notre Service Pour l'Information et la Documentation en Ligne.

Pour interroger les bases de données de SPIDEL, toutes disponibles en français, il vous suffit d'un terminal compatible télétype et d'un téléphone. Vous pouvez ainsi obtenir immédiatement une trace écrite des résultats d'interrogation demandés.

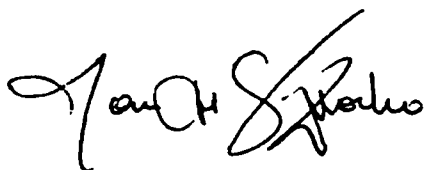
Le fonds documentaire mis à votre disposition par ce processus représente l'analyse d'articles, de rapports, de thèses, de communications de congrès et d'ouvrages divers touchant de nombreux domaines.

Vous pouvez ainsi accéder aux bases de données du CETIM pour la mécanique, du CIS pour l'hygiène et la sécurité du travail, bientôt de l'EDF pour l'électricité et l'électronique. D'autres bases sont prévues pour les prochains mois telles que l'AFEE pour la pollution, PONT-A-MOUSSON pour la métallurgie, CNRS pour l'économie d'énergie (PASCAL), etc.

Si vous n'êtes pas équipés de terminaux et que vous n'envisagez pas de le faire, nous vous conseillons de vous adresser à M. THOLOZAN, Directeur de l'Agence Régionale d'Information Scientifique et Technique (ARIST) de la Chambre Régionale de Commerce et d'Industrie de votre région :  
CRCI Provence Côte d'Azur Corse, BP 287, 13212 MARSEILLES CEDEX 1,  
tél (91) 91 05 00.

Notre service accueil reste en permanence à votre disposition au poste 500 pour répondre à vos questions ou vous mettre en rapport avec les services compétents de la SPI.

Souhaitant collaborer très prochainement avec vous, je reste votre dévoué.



Jean Charles SZABO

P.J. 1 Coupon-réponse.

VOS REF.

NOS REF. LT 115 JCS

CLICHY, le

Madame, Monsieur,

Nous avons le plaisir de vous faire part du lancement de SPIDEL, notre service d'information et de documentation en ligne.

Pour interroger les bases de données de SPIDEL, toutes disponibles en français, il vous suffit d'un terminal compatible télétype et d'un téléphone. Vous pourrez ainsi obtenir immédiatement une trace écrite des résultats d'interrogations demandées.

Le fonds documentaire mis à votre disposition par ce processus représente l'analyse d'articles, de rapports, de thèses, de communications de congrès et d'ouvrages divers touchant de nombreux domaines. (MECANIQUE - THERMIQUE - INFORMATIQUE - SECURITE - PROTECTION - ENERGIES).

Vous pouvez ainsi accéder aux bases de données du CETIM pour la mécanique, du CIS pour l'Hygiène et la Sécurité du Travail, bientôt de l'EDF pour l'électricité et l'électronique. D'autres bases sont prévues pour les prochains mois telles que l'AFEE pour la pollution, Pont à Mousson pour la métallurgie, CNRS pour l'économie d'énergie (PASCAL), etc...

L'utilisateur ne paie aucun abonnement ni droit d'entrée mais uniquement les heures d'interrogations effectuées et les documents édités qu'il a demandés.

Notre Service Accueil reste en permanence à votre disposition au poste 500 pour répondre à vos questions ou vous mettre en rapport avec les services compétents de la SPI.

Souhaitant collaborer très prochainement avec vous, nous vous prions d'agréer, Madame, Monsieur, l'expression de nos sentiments distingués.

Jean Charles SZABO

P.J. : 1 documentation complète SPIDEL.



VOS REF.  
NOS REF. LT 111 AW/JCS

CLICHY, le

Madame, Monsieur,

En réponse à votre demande de renseignements, j'ai le plaisir de vous adresser ci-joint une documentation complète sur SPIDEL, notre Service pour l'Information et la Documentation en ligne.

Pour interroger le fonds documentaire mis à votre disposition, il vous suffit de disposer d'un terminal compatible télétype et d'un téléphone. Vous pourrez ainsi obtenir immédiatement une trace écrite des résultats de vos interrogations. Je peux bien évidemment vous faire une démonstration dans nos locaux si vous le désirez.

L'ensemble des bases de données constituant le fonds documentaire représente l'analyse, par des professionnels, de milliers d'articles, de rapports, de thèses, de comptes rendu de congrès et d'ouvrages divers touchant de nombreux domaines. Je vous remets ci-joint une fiche descriptive pour chacune des bases actuellement opérationnelles.

D'autre part, SPIDEL offre, grâce à la puissance de ses moyens techniques, toutes les facilités de procédures pour effectuer des recherches rapides et précises. L'accès au service se fait à l'aide d'un mot de passe qui vous sera attribué dès la signature du contrat d'utilisation ci-joint.

Vous voudrez bien noter que l'utilisation de SPIDEL n'est assujettie à aucun abonnement ni droit d'entrée, sa seule obligation étant de payer la redevance correspondant au temps d'interrogation et au nombre de documents édités.

Je reste, ainsi que notre service accueil au poste 500, à votre entière disposition pour répondre à vos questions ou vous présenter plus en détail notre service.

Souhaitant vous compter prochainement parmi nos utilisateurs, je reste votre dévoué.

Jean Charles SZABO

COUPON - RÉPONSE

MME MELLE M. : \_\_\_\_\_ FONCTION : \_\_\_\_\_

SOCIÉTÉ : \_\_\_\_\_

ADRESSE : \_\_\_\_\_

TÉL : \_\_\_\_\_

☐ ASSISTERA

☐ N'ASSISTERA PAS

A LA JOURNÉE DE PRÉSENTATION DU

☐ DÉSIRE ÊTRE INFORMÉ DE LA DATE DE LA PROCHAINE SÉANCE

☐ SERA ACCOMPAGNÉ DE : \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

☐ DÉSIRE RECEVOIR UNE DOCUMENTATION SPIDEL

A RETOURNER A : SOCIÉTÉ POUR L'INFORMATIQUE  
Département SPIDEL  
98 Bd Victor Hugo  
92115 CLICHY

SPIDEL