

**UNIVERSITE CLAUDE BERNARD**

UNIVERSITÉ CLAUDE BERNARD LYON I  
D.E.S.S. en INFORMATIQUE DOCUMENTAIRE  
43, Bd du 11 Novembre 1918  
69621 VILLEURBANNE  
Tél. (78) 68.03.27 - P. 416

**PASCAL: SA CONCEPTION ET SA COMPARAISON  
AVEC QUELQUES SYSTEMES DOCUMENTAIRES**

**Note de synthèse présentée par  
Edith BALADI**

**Sous la direction de M. Michel MERLAND,  
Directeur de l'Ecole Nationale Supérieure des Bibliothèques**

**5<sup>eme</sup> promotion 1980**

DESS  
-1980  
J  
A



## SOMMAIRE

INTRODUCTION .....	1
BIBLIOGRAPHIE .....	3
<u>CHAPITRE PREMIER</u>	
I PRESENTATION DE PASCAL .....	7
II ORGANISATION DU FICHER .....	7
I2 SOURCES DE DOCUMENTS .....	8
I2I TYPES DE DOCUMENTS .....	8
I22 REPARTITION DES PERIODIQUES PAR PAYS D'ORIGINE .....	8
I23 REPARTITION DES PERIODIQUES PAR LANGUE .....	8
2 TRAITEMENT DE L'INFORMATION .....	10
2I LE CATALOGAGE .....	10
2II LES DIFFERENTES ZONES DU BORDEREAU .....	10
22 ANALYSE DU CONTENU DU DOCUMENT .....	15
22I CODES DE CLASSEMENT .....	15
222 RESUMES .....	15
223 INDEXATION .....	15
223.1 LE THESAURUS .....	17
223.2 LE LEXIQUE FRANCAIS-ANGLAIS ; ANGLAIS-FRANCAIS .....	17
3 SAISIE .....	17
3I CONTROLE AUTOMATIQUE .....	18
3II CONTROLE DES ZONES CONTROLEES .....	18
3I2 CONTROLE DU TYPE D'INDEX .....	18
3I3 CONTROLE DU CATALOGAGE .....	18
3I4 CONTROLE DE COMPATIBILITE .....	18
3I5 CONTROLE DE FORME .....	18
32 LA CORRECTION EN LIGNE .....	18
4 L'EXPLOITATION DE LA BASE DE DONNEES PASCAL .....	19
4I LE BULLETIN SIGNALETIQUE .....	19
42 LA DIFFUSION SELECTIVE DE L'INFORMATION .....	19
43 LA RECHERCHE RETROSPECTIVE .....	19
44 LES BANDES MAGNETIQUES .....	19

5	SYSTEMES INFORMATIQUES DE RECHERCHE DOCUMENTAIRE .....	20
51	SYSTEME PASCAL VIRA .....	20
52	PASCALINE .....	20

## DEUXIEME PARTIE

INTRODUCTION .....	21
1 IDENTITE DES BIBLIOGRAPHIES .....	28
2 DISCIPLINES COUVERTES PAR LES BASES .....	29
3 NOMBRE ANNUEL DE DOCUMENTS COUVERTS PAR CATEGORIE .....	30
4 NOMBRE TOTAL DE REFERENCES SIGNALES PAR AN.....	30
5 TRAITES COMMUNS DES BIBLIOGRAPHIES ETUDIEES .....	31
6 COMPARAISON ENTRE PASCAL, INSPEC ET COMPENDEX .....	32
61 Comparaison des résultats des 2 recherches: manuelle et automatisée ..	33
62 Le recouvrement .....	37
63 Comparaison des méthodes de sélection et d'indexation par PASCAL, INSPEC et COMPENDEX des articles de quelques revues signalées simul- tanément par les 3 bases .....	37
64 Conclusion .....	38
65 Couverture de la documentation par PASCAL et INSPEC .....	39
7 COMPARAISON ENTRE PASCAL, BIOSIS ET MEDLARS .....	40
71 Etude comparée des bibliographies imprimées: Le Bulletin Signalétique, les Biological Abstracts et l'Index Medicus.....	41
72 Evaluation de la recherche automatique dans PASCAL et MEDLARS .....	41
73 Stratégie de recherche dans le système MEDLARS .....	42
74 Evaluation du rappel dans PASCAL et MEDLARS .....	43
75 Conclusion .....	43
76 Couverture de la documentation par PASCAL, MEDLARS et BIOSIS.....	44
8 COMPARAISON ENTRE PASCAL ET CA-COND .....	45
81 Evaluation du nombre de sorties, de la pertinence et du bruit .....	45
811 Etude du nombre de sorties .....	45
812 Evaluation de la pertinence.....	47
813 Calcul du bruit .....	47
82 Langage d'interrogation .....	47
83 Conclusion .....	47
84 Couverture de la documentation par PASCAL et CA-COND .....	48
9 EVALUATION DES DIFFERENTES BASES PAR L'AGENCE SPATIALE EUROPEENNE .....	50
10 CONCLUSION GENERALE .....	51
ANNEXES .....	
BIBLIOGRAPHIE .....	83

## INTRODUCTION

Une des premières initiatives prises en France, pour la mise en ordre des connaissances revient au Centre de Documentation Scientifique et Technique du CNRS. Créé en 1939 il avait pour mission principale de rassembler les revues scientifiques publiées dans différents pays du monde et d'en faire le signalement aux chercheurs. C'est ainsi que fut créée une bibliothèque de périodiques scientifiques et que fut diffusé, dès 1961 un bulletin bibliographique qui s'appelait alors "Bulletin Analytique" et qui, par la suite a pris le nom de "Bulletin Signalétique". L' "explosion documentaire" survenue depuis 1940 a conduit le CDST à envisager dès 1968 l'information de son fichier documentaire; le projet s'est concrétisé en 1971 avec le Programme Appliqué à la Sélection et à la Compilation Automatiques de la Littérature dit PASCAL .

Pluridisciplinaire, la base de données PASCAL signale un nombre important de documents d'origine très différente et ses fichiers s'accroissent rapidement.

Dans une étude détaillée des techniques appliquées dans la conception de cette base dans toutes ses étapes nous a paru intéressante et nous a poussé à chercher tout ce qui a été écrit sur PASCAL depuis l'année 1974 jusqu'à présent. Cette recherche bibliographique retrospective a montré l'existence d'environ 30 articles dont quelques uns faisaient partie de colloques. Ces articles, bien qu'ils soient écrits dans des revues très différentes sont des articles publicitaires qui, pour la plupart, s'adressent à des usagers n'ayant pas de formation de documentaliste. Ils font une présentation rapide de la base indiquant le nombre de périodiques dépouillés, leur origine... avec ses différents produits et les conditions d'accès. Il est bien regrettable qu'une base aussi importante n'ait pas fait l'objet de publications objectives décrivant d'une façon assez détaillée: les techniques de sa constitution, les difficultés rencontrés lors de sa mise en place , les problèmes auxquels le système est confronté et enfin les lacunes (s'il en existe). De telles publications auraient été d'une grande utilité aux documentalistes actuels cherchant à automatiser leurs services et se trouvant dans une situation transitoire où le passage des systèmes manuels aux systèmes automatisés est devenu d'une grande urgence.

Ce travail est divisé en deux parties : la première consiste à présenter une bibliographie aussi complète que possible sur la base de données PASCAL et à en faire ensuite la synthèse.

Le plan suivi est de décrire dans l'ordre le côté technique de la "chaîne" PASCAL : la réception des documents, les différentes étapes de leur traitement, la saisie et enfin l'exploitation de ses différents produits.

Quant à la deuxième partie elle comprend une tentative de comparaison entre PASCAL et quelques bases de données internationales. Cette comparaison prendra en compte la couverture de la documentation dans chaque base, l'origine des périodiques qu'elle dépouille et la pertinence des documents sélectionnés lors de l'interrogation.

Liste des documents traitant de PASCAL et recensés lors de la recherche manuelle dans la section IOI du Bulletin Signalétique. Cette bibliographie est classée par l'ordre chronologique des volumes du Bulletin.

1974

- 1 D'OLIER (J. H.). DUSOULIER (N.)  
P.A.S.C.A.L. Un système français d'information au service de la médecine.  
Triangle, Fr., (1973), 13, n°3. 235 - 246 .
- 2 BUFFET (P.)  
The PASCAL system  
EUR - 5023 - e , Luxembourg. ( 1973 ), IO2 - IO7.  
( Proc. EUSIDIC conf. Better serv. user ; Luxembourg, 1972 ) .
- 3 DUSOULIER (N.)  
Le Centre de documentation du CNRS.  
Afinidad, Esp., (1973), 30, n° 307, 530 - 532. (Coloq. hisp.- fr.inf. doc. scient. tec. base innovacion teehnol.; Barcelona; 1972 ).
- 4 DUSOULIER (N.) GMACH (M.)  
Thésaurus informatique.  
DGRST - 7I72906 Fr. (1974) IO7 p.1. (Rapp. final, action concertée ; doc. sci. tech.) Iere Ed. CNRS.
- 5 VIGNE (H.)  
Un logiciel bien adapté à l'information scientifique: le système PASCAL-VIRA.  
In: Inf. et doc. 4eme Colloq. Fr. - pol.; Paris; 1974, Paris, A.N.R.T. (1974), I-I7.
- 6 BUFFET (P.)  
P.A.S.C.A.L. 3  
In: Inf. et doc. 4eme Colloq. Fr.-pol. ; Paris; 1974, Paris, A.N.R.T. , (1974), I-2 (Résumé).

1975

- 7 DUSOULIER (N.)  
Bulletin Signalétique du CNRS - PASCAL 3 and Bulletin Signalétique.  
In: Inter. coun. sci. unions abstr. board. proc. gen. assem. meet.; Berlin; 1974, S.I. ICSU AB (s.d.) 99-IOI.

1976

- 8 D'OLIER (J.H.)  
Système PASCAL  
In: Selekt. dystryb. inf. krajowym syst. inte. Mater. Krajowa konf. ;  
Wroclaw ; 1974, Wroclaw, Wyday, politech., (1975) 79-87, res. anglais.  
pol. russe.
- 9 VIGNE (H.) (CNRS, cent. doc. Paris).  
Le service SDI du Centre National de la Recherche Scientifique .  
In: Selekt. dystryb. inf. krajowin syst. inte. Mater. Krajowa konf. ;  
Wroclaw; 1974, Wroclaw, Wyday. polytech., (1975), 109-124, res. pol.  
angl. russe.
- 10 Visite du centre de documentation du CNRS le 10 septembre 1975.  
dans: Inform. et Document., Fr., (1975) n° 3 , 3-8.

1977

- 11 BENASTEAU D. M.  
L'autre mission du CNRS : l'information.  
dans: l'Usine Nouvelle Fr. (1976) n° 33-34, 25
- 12 Voyage à Paris.  
Ass. Bibliothécaires suisses, Ass. suisse Document., Nouv., Suisse, (1977),  
53, n°1, 26-35, res. allem.
- 13 HOFFMANN (W.D.), KLAWITER - POMMER (H.T.)  
Relevante Zeitschriften fur die Veterinarmédizin - Versuch einir Analyse  
anhand von funf Literatur - Datenbasen. (périodiques pertinents pour la  
médecine vétérinaire. Essai d'une analyse à l'aide de 5 bases de données.)  
In: Nachr. Dokument., Dtsch., (1977), 28, n° 2, 64-67, res. angl., bibl. (Bref.)
- 14 PASCAL - VIRA : Manuel d'utilisation. - Centre de Documentation Informascience.-  
Paris, 1977.

1978

- I5 BUHR J. M. & DEGEN C.  
 PASCAL : Une base de données multidisciplinaire, son utilisation en physique atomique et moléculaire et physique des fluides et des plasmas.  
 dans: Journal de physique, Fr. (1977) , 38, n° 3 p. 249-251.
- I6 DUSOULIER Nathalie & BAUD Françoise.  
 Le centre de documentation scientifique et technique du CNRS.  
 dans: Le courrier du CNRS, Fr. (1977) n° 26 p. 46-51.
- I7 BUFFET (P.)  
 PASCALINE, a multidisciplinary European data base.  
 In: On-line inf. Int. on-line inf. I. Meet.; London; 1977, Oxford, Learned information, (s.d.), 143-149.
- I8 On line systems reviewed at international conference.  
 In: Inform. Retrieval Library Automat., U.S.A., (1978), 13, n°8, 1-12.
- I9 D'OLIER (J.H.)  
 Documentation automatisée au CNRS en sciences et techniques.  
 Bull. O.M. Fr., (1977), n° 66, 13-18.
- 20 POISSON (C.) En slovène  
 (Possibilités de recherches bibliographiques en mode conversationnel.  
 Présentation du fichier Pascal pour ce qui concerne le soudage.)  
 In: Varilna Tech., Jugosl., (1977), 26, n° 3, 54-58.
- 21 D'OLIER Jacques & BEUCHE Michèle  
 L'innovation et la documentation automatique.  
 dans: Sciences et techniques, Fr. (1978), n° 50 p. 28-30
- 22 SECHTER (K.)  
 Les méthodes de recherche bibliographique en physiologie humaine. Etude critique de quelques systèmes documentaires.  
 S.l., (1977), 85p., bibl. (5ref.) (Mem. INTD),

- 23 DE JONG - HOFFMAN (M.W.)  
Evaluation of a computer - aided literature search in COMPENDEX, INSPEC  
and PASCAL from the ESA data base concerning the subject: " The application  
of heat pumps for the heating of houses".  
S.l., (1978) (40p.), Trad. du néerl.

1979

- 24 D'OLIER Jacques  
PASCAL: Les banques d'information documentaire du CNRS  
dans: Colloque franco québécois sur l'information scientifique et technique  
de langue française /1978/ Montréal P.Q., CAN, Montréal: Informatech France -  
Québec (1978), 65 - 79.
- 25 BAUD (F.)  
Le Centre de documentation Scientifique et Technique.  
Rev. administrative 1979, vol. 32 n° 188 p. 197 - 201
- 26 EL - MORR Issa  
Traitement et exploitation de l'information.  
Paris: CNRS . 1979. 74p.

## I PRESENTATION DE PASCAL

La base de données PASCAL comporte 3 millions de références sur bandes magnétiques et s'accroît de 500 000 références par an.

Les premières bandes magnétiques exploitables datent de 1972. Les versions PASCAL 1 et PASCAL 2 étaient alors orientées vers la fabrication des différentes sections du Bulletin Signalétique (une bande par section). La version actuelle PASCAL 3, mise en place sur la totalité du fichier depuis 1977 permet une exploitation facile du fichier sous différentes formes:

- Edition du Bulletin Signalétique
- Service de diffusion sélective de l'information
- Recherche rétrospective.

### I I ORGANISATION DU FICHER

Pour traiter ce volume important de documents abordant des domaines très divers, le fichier a été structuré en 9 sous - fichiers (ou unités documentaires) regroupant soit des sciences fondamentales, soit des secteurs d'intérêt.

- U0 Métallurgie
- U1 Biologie végétale. Sciences agricoles. Produits alimentaires.
- U2 Sciences de la terre
- U3 Physique. Informatique. Electricité
- U4 Physique Chimie
- U5 Sciences de la Vie
- U6 Chimie pure et appliquée
- U7 Energie
- U8 Sciences de l'ingénieur

Chaque Unité documentaire comprend un certain nombre de sections scientifiques à l'intérieur desquelles les signalements sont répartis selon un plan de classement.

Des duplications entre sous - fichiers existent. Un même signalement peut être classé dans une ou plusieurs sections du Bulletin Signalétique.

De nombreux organismes participent à la constitution du fichier PASCAL. Le Bureau de Recherches géologiques et minières, le Centre d'Etude de matières plastiques, l'Institut de Soudure, l'Institut National de la Recherche Agronomique sont les plus importants.

## I 2 SOURCES DES DOCUMENTS

Les signalements entrant dans le fichier proviennent de différentes sources primaires, la plus importante d'entre elles étant constituée de périodiques.

### I 2 I Types de documents

- Périodiques: 17 000 titres dont 13 000 courants. L'examen de ces périodiques entraîne l'analyse d'environ 450 000 articles par an (93% du fichier). PASCAL analyse un grand nombre de périodiques soviétiques et des pays de l'Est et un certain nombre de périodiques japonais et latino américains.
- Comptes rendus de congrès : environ 1500 congrès sont traités tous les ans, ce qui correspond à 14 000 signalements (3% du fichier).
- Brevets : 10 000 brevets déposés en France sont dépouillés annuellement dans les domaines techniques (2%).
- Thèses: les thèses françaises de Doctorat d'Etat et d'ingénieur - docteur avec un nombre toujours croissant de thèses de 3eme cycle et de thèses étrangères ce qui porte à environ 7 000 le nombre de thèses signalées par an (1,2%).
- Livres : environ 1 000 ouvrages sont signalés par an (0,6%).
- Rapports techniques: 2 200 rapports par an.(0,4%).

### I 2 2 Répartition des périodiques par pays d'origine

Europe	62%
Amérique du Nord	28%
Asie	5,3%
Amérique latine	4,7%
Afrique, Océanie	

**I 2 3 Répartition des périodiques par langues**

Anglais	57%
Français	12%
Allemand	8%
Russe	12%
Autres (dont Japonais)	11%

## 2 TRAITEMENT DE L'INFORMATION (voir annexe)

C'est à la rédaction que se fait le traitement de l'information. Ceci consiste à faire le catalogage, l'indexation et éventuellement le résumé des documents reçus afin de pouvoir les introduire dans le fichier.

### 2 I LE CATALOGAGE

Pour faciliter le travail et économiser au maximum le temps de saisie, des bordereaux PASCAL ont été élaborés. Ceux ci sont de couleurs selon le type de documents:

- Verte: pour les périodiques
- Bleue: pour les ouvrages, rapports, thèses, normes et miscellaires
- Rouge: pour les brevets
- Orange: pour les traductions
- Violette: pour les rapports DGRST;

et comprenant plusieurs zones et sous zones.

Le catalogage consiste à remplir ces zones selon les règles très précises du Manuel de Référence de l'UNISIST ICSU-AB préconisées par l'UNESCO.

### 2 I I LES DIFFERENTES ZONES DU BORDEREAU

Le nombre des zones et leur nature est en fonction du type du document. Sur le bordereau, la zone est symbolisée par un triangle blanc () appelé drapeau blanc, suivi d'un chiffre ou d'une lettre. La sous-zone est symbolisée par un drapeau noir () suivi d'un chiffre ou d'une lettre.

Toutes les zones ne sont pas necessairement remplies mais il existe des zones obligatoires qui ne peuvent pas être vides, comme il existe des zones à format fixe. Ces deux types de zones sont contrôlés par l'ordinateur parce qu'elles sont des zones de contrôle.

Indications de chaque zone:

-  I Code de la semaine de prise en compte du périodique par le fichier PASCAL. II est composé de trois chiffres; le dernier chiffre de l'année en cours, puis le n° de la semaine.

- ▷ 2 Code représentant le classement (la côte) du périodique à la bibliothèque du CNRS. Alphanumérique, il est sans indice ni exposant.
- ▷ 3 Code de la langue du document analysé, à trois caractères alphabétiques majuscules. Si le document est en plusieurs langues, les différents codes seront séparés par une croix de Lorraine † .
- ▷ 4 Sigle de l'organisme, autre que le CNRS, détenteur du document suivi de la côte du document (celle de l'organisme détenteur) séparée du sigle par un tiret.
- ▷ 5 Code représentant l'organisme, autre que le CNRS, qui a réalisé l'analyse.
- ▷ 6 Code du ou des destinataires de l'analyse.
- ▷ 7 Type bibliographique principal du document:
- |                |                             |
|----------------|-----------------------------|
| P = périodique | S = miscellaire             |
| L = livre      | B = brevet                  |
| R = rapport    | C = compte rendu de congrès |
| T = thèse      | N = norme                   |
- ▷ 8 Code indiquant le type complémentaire auquel peut appartenir le document. Mêmes indices que la zone précédente.
- Exemple:
- ▷ 7P ▷ 8C. cela veut dire qu'on analyse un congrès publié dans un périodique.
- ▷ 9 Code du niveau bibliographique du document. Il est alphabétique à un caractère : A M C .
- A = analytique, indique que le document fait partie d'une entité physique plus large.
- M = monographique où le document forme une entité et ne comporte qu'un seul titre.
- C = collectif pour les documents en plusieurs volumes.

- ▷ M Originalité du document, fonction, style:
- A = archives originales
  - K = atlas
  - Z = bibliographie
  - C = compte rendu
  - L = dictionnaire
  - N = données numériques
  - W = norme
  - Y = "progress report"
  - M = publicité commerciale
  - E = résumé
  - S = synopsis
  - T = tiré à part
  - V = traduction
- ▷ V N° du bordereau pris en compte dans le fichier PASCAL.
- ▷ A Indication du nom de tous les auteurs principaux et secondaires (personnes physiques), dans l'ordre adopté par le document primaire et dans la langue du document. Les noms en caractères cyrilliques sont translittérés, ceux en caractères ni latins ni cyrilliques sont indiqués tel qu'ils se trouvent en caractères latins à la table des matières ou dans le résumé (cf. norme ISO 833).
- ▷ C Indication du nom de tous les auteurs personnes morales (collectivité auteur) . Ce nom figure en entier , en majuscule.
- ▷ O Affiliation ou adresse de l'auteur, elle est toujours composée de 3 segments:
- nom de l'organisation et ses subdivisions.
  - nom de la ville avec éventuellement la division administrative, le code postal.
  - code du pays en trois lettres suivant les normes ISO.

- ▷ E Titre et sous-titre (s) en entier, en minuscule et sans abréviation. Deux cas sont à considérer:
- . Titres écrits dans une langue à caractères latins: dans ce cas le titre original complet est donné dans la langue de la contribution elle-même.
  - . Titre à caractères cyrilliques, chinois etc... dans ce cas le titre original apparait sous une forme translittérée, seulement pour les documents traités au niveau collectif et monographique.
- ▷ T Traduction du titre du document (les titres anglais ne sont pas traduits).
- ▷ F sous zone introduit le titre traduit en français.
- ▷ Q Caractéristiques bibliographiques du documents
- ▷ I : titre abrégé du périodique
  - ▷ 2 : titre complet du périodique
  - ▷ B : n° ISSN
  - ▷ O : code du pays de publication du document
  - ▷ M : date de publication du document
  - ▷ V : tomainson en chiffres arabes
  - ▷ N : N° du fascicule
  - ▷ H : indication d'un n° particulier (n° spécial, supplément)
  - ▷ P : pagination
  - ▷ R : pagination effective : sans publicité avec les bis et les ter; indiquée entre crochets.
  - ▷ F : nombre de hors texte ( planches, cartes, dépliants) en chiffres arabes.
  - ▷ K : code de la langue du ou des résumés figurant sur le document.
  - ▷ L : indication des références bibliographiques contenues dans le document.
  - ▷ A : nombre d'analyses effectuées sur le texte ( ne concerne que le fichier " Sciences de la Terre" ).

▶ S : sous - zone où on met des indications diverses (existence de diapositives, microfiches, microfilms, bandes magnétiques).

▷ X Indication du type complémentaire, utilisée pour les documents ayant les caractéristiques de plus d'un type bibliographique.  
Ex. : un congrès dans un périodique ou une thèse dans un livre.  
Cette zone ne peut exister sans (▷ 8). Il y a 3 hypothèses possibles:

I : a) ▷ 7 P ▷ 8 C ▷ 9 A / M

→ ▷ X , suivi de ≠ 1, introduit le titre du congrès.

b) ▷ 7 P ▷ 8 C ▷ 9 C

→ ▷ X , suivi de ≠ 3, introduit le titre du congrès.  
(cf. normes AFNOR "congrès").

2 : ▷ 7 P ▷ 8 T ▷ 9 A / M

→ ▷ X , suivi de ≠ Y introduit le titre de la thèse (cf. normes AFNOR).

3 : ▷ 7 P ▷ 8 R ▷ 9 A / M

→ ▷ X , suivi de ≠ 3 ≠ 4, introduit le titre du rapport

▷ R C'est la zone du résumé du document qui doit être écrit en français et signé à la fin du texte s'il est fait par un autre centre que le CNRS.

N.B. Il est à noter que le catalogage des brevets présente plusieurs particularités qui n'ont pas été mentionnées ici (26).

2 2 ANALYSE DU CONTENU DU DOCUMENT

2 2 1 CODES DE CLASSEMENT (▷ D)

Chaque signalement est affecté d'un ou de plusieurs codes de classement suivant qu'il est signalé dans une ou plusieurs sections du fichier PASCAL. Il a aussi pour but de répartir les différents articles dans les sections du Bulletin Signalétique.

Ces codes sont toujours du type:

NNN. A. NN. A. NN. A. N. A

N représente un chiffre

A représente une lettre

222 RESUMES

L'analyse du document consiste à rédiger un résumé signalétique ayant pour objet de décrire brièvement les éléments essentiels du document.

Ce résumé peut être donné en plusieurs langues (ex: Métallurgie), dans ce cas le résumé français est en première position sur la bande magnétique.

Les résumés ne reprénnent, en général, pas les indications données dans le titre et n'existent pas lorsque le titre est suffisamment explicite.

2 2 3 INDEXATION

- Type des descripteurs:

Le bordereau comprend 4 zones d'indexation:

01 - 25 : Descripteurs contrôlés : II existe un lexique pour chaque unité documentaire . Pour ce type de descripteurs, un contrôle informatique permet d'obtenir dans le fichier la forme orthographique standard, telle qu'elle apparait dans le lexique, ce qui évite des pertes d'information à la recherche.

26 - 50 : Identificateurs : Ce sont des descripteurs dont on ne peut pas faire une liste à priori.

exemples : Composés chimiques précis

Noms propres

Sigles .....

produit obtenu, un catalyseur ou un solvant réactionnel.

### 2 2 3. I Le Thésaurus

Il est formé d'une liste alphabétique, d'une liste hiérarchique, d'un index alphabétique de la liste hiérarchique, et d'une liste alphabétique permutée.

a - La liste alphabétique: rassemble tous les descripteurs et non descripteurs du thésaurus, classée par ordre alphabétique et accompagnée de toutes les relations existant entre eux.

b - La liste hiérarchique: tous les descripteurs du thésaurus sont ici placés dans leur contexte hiérarchique. Le sommet de chaque arbre est symbolisé par le nom de la facette, indiqué entre parenthèses. Les décalages vers la droite signifient les changements de niveau.

c - L'index alphabétique de la liste hiérarchique: facilite l'accès à la liste hiérarchique.

d - La liste permutée des descripteurs: reprend par ordre alphabétique tous les mots significatifs des descripteurs simples et composés, et donne pour chacun la liste des descripteurs où ils apparaissent.

### 2 2 3. 2 Le lexique français - anglais, anglais - français

C'est un dictionnaire où seuls les descripteurs contrôlés figurent. A l'un terme français correspond un terme anglais et vice versa. Les synonymes n'y sont pas notés.

Ce lexique est enregistré dans le fichier PASCAL; il permet, d'une part d'interroger ce fichier en français et en anglais et d'autre part de publier, dans la bibliographie imprimée, des index avec une entrée en anglais en plus des index français.

### 3 SAISIE

Les bordereaux sont saisis sur bande perforée qui, à l'aide d'une photocomposeuse, permet d'obtenir un bromure utilisé pour une première correction. Les mots y sont numérotés. Ensuite la correction est portée sur

une bande perforée qui avec la première bande permet d'obtenir une troisième bande corrigée et un bromure. Cette bande est ensuite transformée en bande magnétique qui entre dans la constitution du fichier PASCAL.

### 31 CONTROLE AUTOMATIQUE

Il existe des programmes de contrôle qui détectent et signalent toutes les erreurs qui pourraient avoir lieu dans la rédaction du bordereau ou dans la saisie.

Le contrôle automatique se situe à plusieurs niveaux :

311 Contrôle des zones contrôlées ( ▶ I à ▶ 25) : Ceci consiste à vérifier si chaque terme descripteur existe déjà dans l'index.

312 Contrôle du type d'index : Comme les sections du Bulletin Signalétique sont indexées soit en KWIC soit en KWOC, le contrôle consiste à vérifier si le type d'index dans la zone ▷ Y correspond à celui de la section signalée dans ▷ D.

313 Contrôle du catalogage : Ce contrôle vérifie les séquences des zones; or ces zones doivent être dans un ordre bien précis et quelques unes sont obligatoires.

314 Contrôle de compatibilité entre les différentes zones du bordereau. Une indication dans une zone appelle obligatoirement d'autres indications dans d'autres zones.

315 Contrôle de forme propre aux zones de format fixe et à quelques unes de format libre.

### 32 LA CORRECTION EN LIGNE

Il est possible de faire la correction du fichier PASCAL en ligne par l'intermédiaire d'une console. Pour ceci un programme est mis en place permettant de rectifier le contenu de toutes les zones. Le numéro de saisie ( ▷ V) est la clef de l'enregistrement.

#### 4 L'EXPLOITATION DE LA BASE DE DONNEES PASCAL

##### 41 LE BULLETIN SIGNALÉTIQUE

Réalisé à partir du fichier PASCAL par photocomposition. Le Bulletin comporte 51 sections dont les sujets sont regroupés soit par disciplines soit par centres d'intérêt spécialisés. Les signalements sont rassemblés suivant un plan de classement et peuvent être consultés facilement à l'aide d'index matières et auteurs. Depuis 1972 on peut obtenir les sections du Bulletin sur microfiches.

##### 42 LA DIFFUSION SELECTIVE DE L'INFORMATION

Le service de diffusion selective de l'information sur profils est prévu pour répondre à des besoins documentaires plus ponctuels et à des exigences de rapidité, de souplesse et de pertinence. Les profils sont des bibliographies mensuelles se rapportant à des sujets très précisément délimités par le CDST (profil standard) ou définis par l'utilisateur lui-même (profil personnalisé).

##### 43 LA RECHERCHE RETROSPECTIVE

Consiste à rechercher tout ce qui a été publié sur un sujet déterminé pendant une période donnée. Elle est faite par l'ordinateur pour les années postérieures à 1971 à l'aide du programme PASCAL - VIRA.

##### 44 LES BANDES MAGNETIQUES

Elles représentent tout ou partie de la base de données PASCAL. Elles peuvent être diffusées par sections du bulletin ou correspondre à des profils. Une documentation fournie à la demande décrit toutes les caractéristiques de ces bandes.

## 5 SYSTEMES INFORMATIQUES DE RECHERCHE DOCUMENTAIRE

### 51 SYSTEME PASCAL - VIRA

Le système VIRA est un ensemble de programmes permettant de faire de la recherche documentaire par ordinateur. Ce système est utilisé pour le traitement des profils et des recherches rétrospectives. Le traitement se fait en différé par " lots ".

Sa particularité est de comprendre un système de hashcode qui fait gagner beaucoup de temps et de place en mémoire rapide. Le hashcode est un procédé de codification qui consiste à remplacer chaque mot clé par un numéro. Comme on ne peut pas faire cette opération manuellement (car les risques d'erreur seraient grands) le hashcode consiste à faire générer par les lettres des mots clés eux-mêmes le code numérique qui va les remplacer. Ainsi on peut effectuer plusieurs opérations arithmétiques simples (addition, multiplication...). Moyennant quoi pour un nombre de questions et de documents élevé le temps de traitement peut être très réduit.

### 52 PASCALINE

C'est le système d'exploitation de PASCAL en ligne. L'interrogation de PASCAL en ligne se fait à l'aide d'un terminal qui peut être connecté à l'ordinateur central par un téléphone muni d'un modulateur.

Le fichier PASCAL est implanté sur deux ordinateurs serveurs:

- Sur le système de l'Agence Spaciale Européenne à Frascati où PASCAL est interrogeable par le logiciel RECON.
- Sur Télésystème à Valbonne où PASCAL est interrogeable par le logiciel MISTRAL.

DEUXIEME PARTIE

## INTRODUCTION

L'interrogation des bases de données est devenu une nécessité pour tout travail de recherche. Et, vu le nombre toujours croissant de ces bases avec le coût élevé de leur interrogation, l'utilisateur est obligé d'avoir une certaine idée sur les caractéristiques des différentes bases concernant son domaine de recherche afin de pouvoir les utiliser de la manière la plus efficace.

Ces bases ne cernent pas la documentation scientifique de la même manière. Deux questions se posent à celui qui veut les comparer : jusqu'à quel point ces bibliographies couvrent - elles la littérature? et par quels moyens permettent - elles d'accéder aux références (classifications, thésaurus, logiciels ....)?

Donc la différence d'une bibliographie à l'autre se fait sur des problèmes de couverture et de structure, que ces bibliographies soient étudiées sous leur forme imprimée ou automatisée.

Concernant la comparaison entre différentes bases de données il y a eu peu de travaux menés dans ce sens. Cette question est d'une telle complexité qu'un chercheur isolé ne semble pas en mesure de la résoudre. Il s'agit ici de quelques bases très connues : PASCAL, INSPEC, CA - COND, BIOSIS, MEDLARS et COMPENDEX.

Ce travail consiste à recenser les documents qui se sont plus ou moins intéressés à certains aspects de cette comparaison et à en faire la synthèse. Dans ce but une recherche manuelle a été faite dans le Bulletin Signalétique suivie d'une recherche automatisée dans les deux bases : PASCAL et INSPEC.

L'idée courante étant que les bibliographies spécialisées couvrent le mieux leur domaine, notre approche est une tentative de situer la base PASCAL par rapport aux autres pour chaque spécialité.

Dans cette comparaison nous dégagerons les grands traits communs à ces bibliographies ensuite nous présenterons les résultats obtenus lors des études comparatives antérieures des différentes sections de PASCAL avec les bases correspondantes et enfin nous tenterons de voir dans quelle mesure elles couvrent la documentation et se recoupent.

## BIBLIOGRAPHIE

1975

## 1 LIBES (Michel)

Bibliographies en Métamorphose

dans: Bull. Bibl. France, Paris, vol. 20, n° 8 , 1975

1976

## 2 LARSON (J.R.) &amp; BERNHARD (G.A.)

Comparison of printed bibliographic descriptions distributed by BIOSIS, CAS, EI

in: J. Amer. Soc. Inform. Sci., USA, (1976) 27 n° I p. 46 - 52.

## 3 DURKIN (K.) &amp; SMITH (J.R.)

Retrieval forms and formats for environmental objectives employing machine readable biological information.

in: J. Chem. Inform. Computer Sci., USA (1976), 16, n° I , 14 - 16.

1977

## 4 HOFFMANN (W.D.), KLAWITER - POMMER (H.)

Relevante Zeitschriften für Versuchseinheit Analyse anhand fünf Literatur Datenbase.

in: Nachr. Document. , Dtsch., (1977), 28, n° 2, 64 - 67.

## 5 HAWKINS (D.T.) , MILLER (B.)

On line data base coverage of the on line information - retrieval literature.

On line Rev., G.B., (1977), I, n° I, 59 - 64

## 6 WILLIAMS (M.E.)

Analysis of terminology in various CAS data files as access points for retrieval .

in: J. Chem. Inform. Computer Sci., USA (1977) 17, n° I , 16 - 20.

## 7 SEWELL (W.) &amp; BEVAN (A.)

Nonmediated use of MEDLINE and TOXLINE by pathologists and pharmacists.

in: Bull. med. Library Ass., USA (1976) 64, n° 4, 382 - 391 , bibl.

8 BONNERBE - WOLFF (F.)

De la conduite des essais d'évaluation des systèmes d'information et de leurs langages.

S.L. (1975), Mémoire INTD Paris.

9 SCHENK (H.R.) & WEGMULLER (F.).

Substructure search by means of the Chemical Abstracts service chemical registry 2 system.

in: J. Chem. Inform. Computer Sci., USA (1976) 16, n° 3 153 - 161 bibl.

10 FISHER (D.A.)

Keeping Current through information services.

in: Food Technol., USA, (1976) , 66-68.

11 HUBER (A.L.) & MOLL (W.)

A broken hip and MEDLINE : a case report.

in: Bull. med. Library Ass., USA (1976), 64, n° 3 , 320 - 322.

1978

12 TAGLIACOZZO (R.)

Estimating the satisfaction of information users .

in: Bull. med. Library Ass., USA, (1977), 65, n° 2, 243 - 249 bibl. (8ref.)

13 CITTADINO (M.L.) & GIESE (R.L.) & CASWELL (J.V.)

Three computer based bibliographic retrieval system for scientific literature.

in: Bio Science, USA (1977) , 27, n° II, 739- 742 bibl. (1ref.)

14 FALKENBERG (G.)

Medline at BMDC, Karaliuska Instituted

in: Interact. libr. comput. processes libr. inf. network. Proc. Semelin.; Stockholm; 1974, Stockholm, Swed. soc. tech. doc., (1975) 155 - 173.

- 15 HARLEY (A.J.)  
The UK MEDLARS service : a personal view of its first decade.  
in: Aslib Proc., G.B. , (1977) 29, n° 9, 320 - 325.
- 16 MILLSON (D.R.) , MAYATT (A.G.)  
The Medlars databases and the UK contribution to MeSH indexing .  
in: On line inf. Int. on-line inf. I. Meet.; London; 1977, Oxford,  
learned information, (s.d.) I39-I42.
- 17 BUFFET (P.)  
PASCALINE : a multidisciplinary European data base.  
in: On - line inf. Int. on - line inf. I. Meet; London, 1977, Oxford,  
Learned information I43 - I49.
- 18 On line systems reviewed at international conference.  
in: Infor. Retrieval. Library Automat., USA (1978) I3, n° 8 p. I - I2.
- 19 REMY (D.)  
Les Chemical Abstracts et l'interrogation de ses bandes en conversationnel  
sur le système ESA / RECON.  
dans : Bull. Dica, Fr., (1977), 2 , n° II - I2, 5 - I6.
- 20 REMY (D.)  
Interrogation des Chemical Abstracts sur le système Lockheed / Dialog.  
dans: Bull. Dica, Fr. (1977), 2, n° II - I2, I7 - 32.
- 21 SOLAL (E.)  
Après une recherche analytique sur un même sujet .....international.  
S.L. (1977) I - 75 , bibl. (2 p.) Mémoire INTD.
- 22 Medlearn : A computer - assisted instruction , program for Medlars.  
in: Bull. med. Library Ass., USA (1978), 66, n° I , 6 - I3.
- 23 VAN CAMP (A.J.) & FOREMAN (G.)  
Biosis Previews & Medlars - A biomedical team.  
On line , USA (1977), I, n° I, 24 - 29.

- 24 HAFNER (A.W.), HAYCOCK (L.A.) & CAROLL (D.J.)  
 Searching the MEDLARS spécial list oategories.  
 in: Online , USA (1977) I, n° 3, 73 - 81.
- 25 BASER (K.H.) & COHEN (S.M.) & DAYTON (D.L.)  
 On line indexing experiment at CAS : algorithmic generation of articulated index entries from natural phrases.  
 in: J. Chem. Inform. Computer Sci., USA, (1978) 18, n° 1, 18 - 25.
- 26 DATTA (V.)  
 Coverage of specialised biological information by online data bases and comparision of this coverage with an in house manual system...  
 Program, G.B. (1978), 12, n° 2 , 55 - 63.
- 27 O'DETTE (R.E.)  
 The CAS data base.  
 Pure Appl. Chem., G.B. (1977), 49, n° 12, 1781 - 1792 bibl. (8 ref.).

1979

- 28 BURROWS (S.) & KYLE (S.)  
 Searching the MEDLARS file on NLM and BRS: a comparative study.  
 in: Bull. med. Library Ass. USA , 1979, 67, n° 1 , 15 - 24.
- 29 REVERSE (A.)  
 Le logiciel MISTRAL 3/4.  
 Documentaliste, FRA (1979), 16, n° 1, 15 - 18.
- 30 ELIAS (A.) :  
 Marketing for online bibliographic services.  
 in: Online Rev. GBR, (1979), 3, n° 1, p. 107 - 117 (INTD)

1980

31 MANTEL MESQUI (Maryvonne)

Comparaison des informations fournies par 4 bases de données américaines dans le domaine agronomique (recherche rétrospective)

FRA, (1978) (2) - 101 p. - (3) f. de pl., 30 cm., bibl. (2p.) Mém.

INTD : Sci. Techn. Inform. et Doc. / (Paris) CNAM.

## I IDENTITE DES BIBLIOGRAPHIES CONFRONTEES

<p align="center"><b>Producteur</b></p> <p align="center">Sigle - nom - adresse</p>	<p align="center"><b>Titre de la bibliographie</b></p> <p>a - sous forme imprimée</p> <p>b - appellation du système automatisé</p>
<p>CNRS    CENTRE NATIONAL DE LA RECHER- CHE SCIENTIFIQUE. Centre de Documentation. Paris - France</p>	<p>a - Bulletin Signalétique</p> <p>b - PASOAL</p>
<p>INSPEC    INFORMATION SERVICES IN PHY- SICS, ELECTROTECHNOLOGY, COM- PUTERS &amp; CONTROL . London - Grande Bretagne.</p>	<p>a - Comprend 4 revues d'abstracts</p> <ul style="list-style-type: none"> <li>- Physics Abstracts</li> <li>- Electrical &amp; Engineering Abstracts</li> <li>E Electronic Abstracts</li> <li>- Computer &amp; Control Abstracts.</li> </ul> <p>b - INSPEC</p>
<p>CAS    CHEMICAL ABSTRACTS SERVICE, a Division of the American Chemical Society. Columbus - USA</p>	<p>a - Chemical Abstracts</p> <p>b - plusieurs appellations :</p> <ul style="list-style-type: none"> <li>- CA - COND (Chemical Abstracts Condensates)</li> <li>- CHEMCON , CAC , CA-CONDENS.</li> </ul>
<p>BIOSIS    BIOSCIENCES INFORMATION SER- VICE. Philadelphia - USA</p>	<p>a - Deux revues complémentaires</p> <ul style="list-style-type: none"> <li>- Biological Abstracts (BA)</li> <li>- Bioresearch Index (BIOI)</li> </ul> <p>b - BIOSIS PREVIEWS.</p>
<p>NLM    NATIONAL LIBRARY OF MEDICINE</p>	<p>a - Index Medicus</p>

Bethesda - USA	b - Medical Literature Analysis and Retrieval System. MEDLINE = MEDLARS on line.
EI ENGINEERING INDEX, Inc. New York - USA	a - The Engineering Index b - COMPENDEX = Computerized Engineering Index.

## 2 DISCIPLINES COUVERTES PAR LES BASES

PASCAL : Pluridisciplinaire

INSPEC : Physique  
 Electronique  
 Informatique et commande automatique.

CA - COND: Biochemistry  
 Organic chemistry sections  
 Macromolecular sections  
 Applied chemistry and chemical engineering sections  
 Physical and analytical chemistry sections.

BIOSIS : General biology  
 Basic medical sciences  
 Microbiology, Immunology, Public health and parasitology  
 Plant sciences  
 Animal sciences.

**MEDLARS : Medecine**  
**Biologie**  
**Biochimie**  
**Pharmacologie**

**COMPENDEX : Engineering, Instrumentation dans les domaines:**  
**Chimie, Agriculture, Alimentation, Electricité, Electronique,**  
**Télécommunications, Mécanique, Informatique, Aéronautique, Mine.**

**3 NOMBRE ANNUEL DE DOCUMENTS COUVERTS PAR CATEGORIE**

	PASCAL	INSPEC	CA - COND	BIOSIS	MEDLARS	COMPENDEX
Périodiques	13.000	2.400	14.000	8.000	3.000	3.500
Livres	1.000	500	500	1.200		+
Congrès	1.500	20.000	700	(BA)avec les		+
Brevets	10.000	3.000	70.000	congrès		
Thèses	7.000	2.500	+	+		
Rapports	2.200	4.000	+	+		+

**N.B. Les catégories de documents présents dans la base et dont le nombre annuel n'est pas connu sont désignées par une croix.**

**4 NOMBRE TOTAL DE REFERENCES SIGNALÉES PAR AN**

PASCAL	500.000	BIOSIS	290.000
INSPEC	180.000	MEDLARS	240.000
CA - COND	380.000	COMPENDEX	85.000

## 5 TRAITS COMMUNS DES BIBLIOGRAPHIES ETUDIÉES

Bien que les bibliographies spécialisées actuelles traitent des documents rédigés en toutes langues, les signalements et les analyses sont exprimés en une seule langue. Ces bibliographies ont connu un développement rapide consécutif à celui de la littérature primaire surtout entre les années 1958 et 1973. Naturellement le total général des références publiées par chacune de ces bibliographies ne donnerait qu'une image très floue sur la production scientifique. En effet cette production ne peut être que très partiellement représentée par le Bulletin Signalétique qui ne couvre que 13.000 revues, quant aux bibliographies spécialisées comme INSPEC et les Chemical Abstracts elles sont souvent selectives dans les articles qu'elles dépouillent. Donc aucune bibliographie ne pourrait être exhaustive . D'ailleurs un grand nombre de revues sont dépouillées par plusieurs bibliographies simultanément.

## 6 COMPARAISON ENTRE PASCAL, INSPEC ET COMPENDEX

Dans le domaine de la physique la comparaison entre PASCAL, INSPEC et COMPENDEX a été faite d'après une recherche dans les 3 systèmes sous leurs deux formes: imprimée et automatisée.

Afin de bien réussir la recherche automatisée, une étude a porté sur la manière dans laquelle chaque bibliographie est compilée. Donc pour chacune on a examiné:

- Les caractéristiques et le nombre des références.
- Les domaines couverts et les divisions correspondantes.
- Les procédures d'accès aux signalements que ce soit dans les versions imprimées ou dans les systèmes automatisés.

Cette étude a abouti aux résultats suivants:

- . PASCAL, INSPEC et COMPENDEX couvrent chacune plusieurs disciplines dans le domaine de la physique. Alors que PASCAL passe pour la base qui signale le plus de documents, COMPENDEX est celle qui signale le moins.
- . INSPEC est la plus claire concernant le classement et l'organisation de sa bibliographie. Les compilateurs de COMPENDEX et de PASCAL ne donnent qu'une explication très succincte sur la manière avec laquelle leurs travaux sont compilés.
- . Les modes de classement et d'indexation sont différents d'une bibliographie à l'autre. Dans chacune des trois bibliographies les indexeurs donnent à chaque signalement au moins 1 code de classement. Ce code couvre un domaine assez vaste dans INSPEC et COMPENDEX, tandis que dans PASCAL il est si restrictif à tel enseigne qu'il correspond à un mot matière (descripteur); autrement dit le classement systématique des signalements dans PASCAL est le plus fin.
- . Dans les 3 bibliographies l'indexation se fait d'après un vocabulaire contrôlé pris dans un thésaurus. Mais dans INSPEC il y a en plus une

indexation par un vocabulaire libre (Free Terms) : ce sont les mots importants du titre. Donc la recherche par les termes du thésaurus se fait sur le champ mots clés des signalements dans les 3 bibliographies. La recherche en langage libre se fait sur le champ "Free Terms" dans INSPEC et sur les mots du titre dans PASCAL et COMPENDEX.

- INSPEC publie régulièrement son thésaurus avec ses mises à jour tandis que PASCAL ne publie qu'une partie de son thésaurus .

Après cette étude une recherche a été effectuée dans PASCAL, INSPEC et COMPENDEX sur le sujet suivant:

" The application of heat pumps for the heating of houses".

Cette recherche a été faite manuellement dans les versions imprimées et en ligne à l'aide d'un terminal.

Dans la recherche en ligne plusieurs stratégies ont été employées. Or la question a été posée à chaque base sous 5 formes différentes correspondant à 5 combinaisons différentes de mots clés.

Ces combinaisons sont les suivantes:

- A. Heat pumps and heating and houses not sun energy:  $1 \times 2 \times 3 - 4$  .
- B. Heat pumps and space heating not sun energy:  $1 \times 2 \times 3 - 4$  .
- C. Heat pumps and heating not sun energy :  $1 \times 2 - 4$  .
- D. Heat pumps and houses not sun energy :  $1 \times 3 - 4$  .
- E. Heat pumps and ( heating or houses or space heating ) not sun energy:  
 $1 \times (2 + 3 + 3 A) - 4$  .

#### 6I Comparaison des résultats des deux recherches : manuelles et automatisée

Le tableau ( ) montre le nombre de références obtenues dans chaque base.

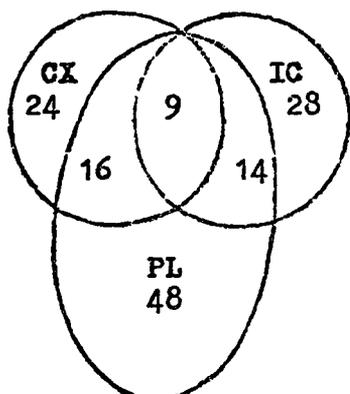
D'après la recherche automatisée, la moyenne des références pertinentes recensées par PASCAL a été 43,8 ref., celle de COMPENDEX 21 ref. alors que INSPEC n'a recensé que 16,2 ref.

Il est clair que PASCAL a donné le plus grand nombre de de références pertinentes dans les deux recherches : manuelle et automatisée.

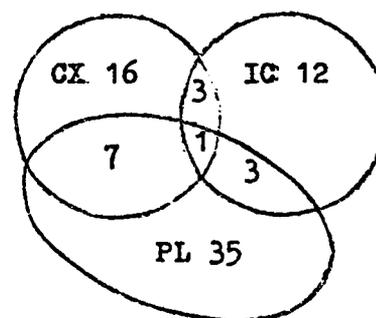
Tableau I : Nombre de signalements pertinents retrouvés

In	Manual	Computer-aided search with combination				
		A	B	C	D	E
<b>COMPENDEX</b>						
exclusive in this	8	7	13	8	5	6
also in INSPEC or PASCAL	7	8	9	12	12	15
also in INSPEC and PASCAL	9	1	1	3	1	4
<b>Total</b>	<b>24</b>	<b>16</b>	<b>23</b>	<b>23</b>	<b>18</b>	<b>25</b>
<b>INSPEC</b>						
exclusive in this	14	7	1	7	11	9
also in COMPENDEX or PASCAL	5	4	3	7	9	13
also in COMPENDEX and PASCAL	9	1	1	3	1	4
<b>Total</b>	<b>28</b>	<b>12</b>	<b>5</b>	<b>17</b>	<b>21</b>	<b>26</b>
<b>PASCAL</b>						
exclusive in this	27	26	28	40	27	38
also in COMPENDEX or INSPEC	12	8	8	11	9	14
also in COMPENDEX and INSPEC	9	1	1	3	1	4
<b>Total</b>	<b>48</b>	<b>35</b>	<b>37</b>	<b>54</b>	<b>37</b>	<b>56</b>
<b>COMPENDEX + INSPEC + PASCAL</b>						
retrieved 1 time	49	40	42	55	43	53
retrieved 2 times	12	10	10	15	15	21
retrieved 3 times	9	1	1	3	1	4
<b>Total</b>	<b>70</b>	<b>51</b>	<b>53</b>	<b>73</b>	<b>59</b>	<b>78</b>

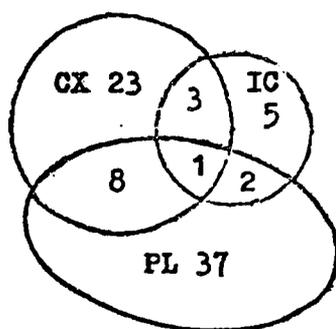
Tableau 2 : Couverture des différentes bases selon les méthodes de recherche



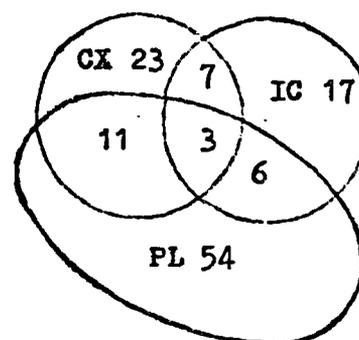
Manual  
retrieved 70 relevant abstracts



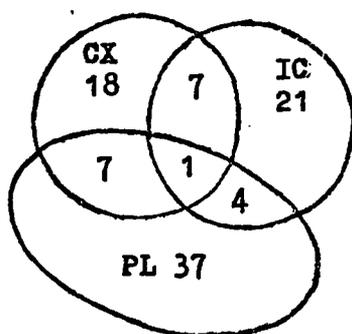
Computer-aided: Combination A  
retrieved 51 relevant abstracts



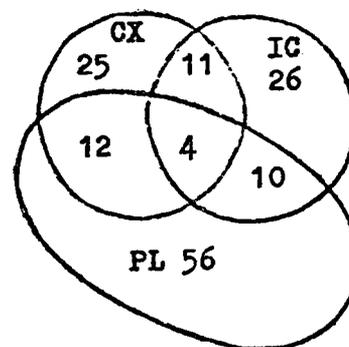
Computer-aided: Combination B  
retrieved 53 relevant abstracts



Computer-aided: Combination C  
retrieved 73 relevant abstracts



Computer-aided: Combination D  
retrieved 59 relevant abstracts



Computer-aided: Combination E  
retrieved 78 relevant abstracts

Tableau 3 : Comparaison de l'efficacité des deux recherches: manuelle et automatisée

a. Name of the bibliography	COMPENDEX					INSPEC					PASCAL					COMPENDEX + INSPEC + PASCAL				
b. Total number of abstracts retrieved, manually + on-line	27					34					65					88				
c. Number of abstracts retrieved manually	24					28					48					70				
d. Percentage c./b.	89					82					74					80				
-----																				
e. Applied combination by the computer-aided search	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
f. Number of classification codes and keywords that were used	16	8	7	15	17	17	6	6	16	18	10	6	6	9	11	28	14	13	26	31
-----																				
g. Number of abstracts in the formed set	13	78	80	16	83	15	40	62	24	71	41	74	177	47	183	69	192	319	87	337
h. Number of relevant abstracts there from	12	18	18	14	20	12	5	17	21	26	35	37	58	39	60	59	60	93	74	106
i. Percentage h./g.	92	23	23	88	24	80	13	27	88	37	85	50	33	83	33	86	31	29	83	31
-----																				
j. Number of by h. retrieved references	16	23	23	18	25	12	5	17	21	26	35	37	54	37	56	51	53	73	59	78
k. Percentage j./b.	59	85	85	67	93	35	15	50	62	77	53	56	83	56	86	58	62	82	78	88
-----																				
l. Percentage j./88 (maximum)	18	26	26	20	28	14	6	19	24	30	40	42	62	42	64	58	62	82	78	88



Bien que le sujet de recherche soit simple il est difficile de préciser quelle est la meilleure stratégie de recherche. Les trois bibliographies ont contribué différemment au recensement des documents pertinents et chacune a ses points forts. Le jugement porté sur la performance d'une bibliographie d'après un seul sujet n'est pas très prudent.

## 62 Le recouvrement

C'est le nombre de références identiques trouvées dans les trois bases. II varie selon le type de recherche (voir tableau n° ).

## 63 Comparaison des méthodes de sélection et d'indexation par PASCAL,

INSPEC et COMPENDEX des articles de quelques revues signalées simultanément par les 3 bases. Pour cette recherche il était important de trouver le plus de littérature possible concernant le sujet . Pour cela on a dépouillé tous les volumes du " Thermal Abstracts " compris entre 1974 et 1977 ( c'est la même période sur laquelle s'est étendue la recherche dans les 3 bases). Cet outil de référence recense exclusivement les sujets concernant les installations dans les batiments. Sa consultation a donné 107 articles tirés de 46 revues différentes. Il s'est avéré que de ces 46 revues 21 sont également dépouillées aussi bien par COMPENDEX que par INSPEC, alors que PASCAL en dépouille 31. Neuf de ces 46 revues sont dépouillées simultanément par chacune des 3 bases et contiennent 31 articles concernant le sujet.

D'après la recherche de ces articles dans les 3 bases on a abouti à la conclusion suivante:

- Le plus grand nombre de ces articles a été trouvé dans PASCAL.
- Dans COMPENDEX la stratégie de recherche a beaucoup moins d'influence sur le nombre de documents retrouvés que dans PASCAL et dans INSPEC.
- La recherche avec un seul profil dans plusieurs bases donne un meilleur résultat que la recherche avec plusieurs profils dans une seule base.

Ces différences d'une base à l'autre sont déterminées par les facteurs suivants:

- a - Selection des articles par les analystes.
- b - Choix des descripteurs affectés aux documents par les indexeurs.
- c - Les index matière mis en mémoire d'ordinateur sont parfois incomplets (ex. PASCAL)
- d - Les titres et les résumés d'auteurs qui reflètent le contenu des articles ne sont pas souvent explicites.

Aucune des trois bibliographies ne signale tous les articles des revues qu'elle dépouille.

- . Les critères de choix des articles qui doivent être signalés sont subjectifs et varient d'un indexeur à l'autre. Mais ce qui est évident c'est que COMPENDEX opère une plus forte sélection que INSPEC et PASCAL.
- . Il y a une grande différence dans les termes des thésaurus disponibles pour ce sujet de recherche. COMPENDEX a des termes très précis, ceux d'INSPEC et de PASCAL le sont beaucoup moins.
- . Les termes disponibles du thésaurus ne sont pas toujours bien affectés aux articles indexés. Dans cet exemple PASCAL a donné le meilleur résultat (précision 81%).
- . L'indexation dans INSPEC par les termes contrôlés et non contrôlés était insuffisante. L'indexation dans PASCAL par les termes contrôlés était à peu près parfait cependant le manque de termes libres indispensables pour une bonne recherche a donné des résultats déficients.

#### 64 Conclusion

COMPENDEX, à cause du nombre limité des revues qu'elle dépouille et de son caractère trop sélectif a recensé peu d'articles sur le sujet étudié. Mais ces articles ont été faciles à retrouver.

INSPEC a recensé un nombre d'articles plutôt élevé à cause du grand nombre de ses sources et de sa souplesse de sélection. De par sa

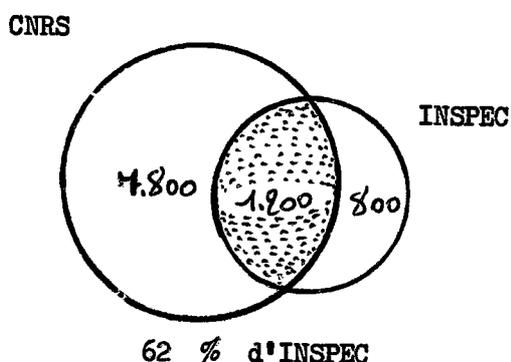
mauvaise indexation il a été difficile de trouver les articles pertinents lors de l'interrogation.

PASCAL a eu une bonne contribution dans les documents recensés. Tout comme INSPEC elle se caractérise par sa richesse en revues sources et sa souplesse dans la sélection.

#### 65 Couverture de la documentation par PASCAL et INSPEC

PASCAL assure une bonne couverture générale de la littérature publiée dans le domaine de la physique quoiqu'elle ne traite que près de 62% des revues analysées par INSPEC (d'après une étude faite en 1975). Or les revues de base qui fournissent la majeure partie des références en physique sont toutes dépouillées par le Bulletin Signalétique. Actuellement INSPEC traite près de 800 revues non dépouillées par le Bulletin Signalétique. Mais en vertu de sa pluridisciplinarité le Bulletin traite un grand nombre de revues qui contiennent beaucoup d'articles intéressant la physique sans être vraiment spécialisées en physique.

Il est à noter que INSPEC a une meilleure représentation des revues japonaises que PASCAL, et aussi une plus large couverture de revues d'électricité et d'électronique.



## 7 COMPARAISON ENTRE PASCAL, BIOSIS ET MEDLARS

### 7I Etude comparée des 3 bibliographies imprimées: Le Bulletin Signalétique, les Biological Abstracts et l'Index Medicus

Une recherche rétrospective manuelle a été effectuée sur le sujet suivant: "Prolactine au cours du cycle menstruel et de la grossesse" portant sur l'année 1976 à l'aide des bibliographies imprimées publiées par les 3 systèmes.

Cette recherche a permis de dégager les différences entre les 3 bibliographies:

- A la différence de l'Index Medicus qui couvre l'aspect clinique et expérimental de la biomédecine, le Bulletin Signalétique (Unité documentaire Sciences de la Vie U5) et les Biological Abstracts s'attachent spécialement à son aspect fondamental.
- Dans le Bulletin Signalétique et les Biological Abstracts les notices sont classées systématiquement et précédées de leurs numéros d'identification; quant aux index, ils présentent les mots clés suivis des numéros des références correspondants. Dans l'Index Medicus les mots sont classés alphabétiquement et les références bibliographiques complètes des documents sont indiquées directement à leur suite.
- A part les Biological Abstracts qui utilisent le langage naturel, il existe un dictionnaire de termes normalisés constituant un simple lexique pour le Bulletin Signalétique et présenté sous forme de thésaurus élaboré pour l'Index Medicus.
- Les Biological Abstracts, malgré leur accessibilité en langage naturel sont de consultation difficile en raison de la complexité des index.
- L'Index Medicus utilise un vocabulaire qui peut paraître trop limité à des termes assez généraux mais il a l'avantage de proposer au lecteur un thésaurus, le MESH indiquant les renvois.

- Les sections de l'Unité Sciences de la Vie (U5) du Bulletin Signalétique utilisent un vocabulaire plus riche mais celui-ci n'est exposé que dans le lexique correspondant (lexique U5) sans aucune indication des termes éliminés.
- Les notices bibliographiques de l'Index Medicus ne comportent pas de résumés.
- Les résumés dans les Biological Abstracts sont plus longues que celles du Bulletin Signalétique.

**72 Evaluation de la recherche automatique dans PASCAL et MEDLARS**

Pour juger les performances de la recherche automatisée dans le système PASCAL on a calculé la précision et le rappel correspondant à la recherche sur " Prolactine au cours du cycle menstruel et de la grossesse", effectuée dans les 2 systèmes.

Si on considère:

- A= nombre de documents pertinents sélectionnés
- B= nombre de documents parasites sélectionnés
- C= nombre de documents pertinents non sélectionnés
- D= nombre de documents parasites non sélectionnés

Le rappel sera:

$$\text{Rappel} = \frac{A}{A + C}$$

et la

$$\text{Précision} = \frac{A}{A + B}$$

L'interrogation de chacune des 2 bases a porté sur les fichiers des deux années : 1976 et 1977.

Les résultats obtenus ont été les suivants:

- La valeur de la précision pour l'année 1976 a été de :
- 58,8 % dans PASCAL
- 48,2 % dans MEDLARS

La valeur de la précision pour l'année 1977 :

74,2 % dans PASCAL

57,1 % dans MEDLARS.

Donc pour ce sujet on a obtenu une meilleure précision dans le système PASCAL que dans le système MEDLARS, sans oublier que la subjectivité intervient souvent dans le jugement de la pertinence.

Il est à noter que dans le système MEDLARS la pertinence est plus difficile à déterminer car les références ne contiennent pas toujours un résumé.

### 73 Stratégie de recherche dans le système MEDLARS

Comme dans le système PASCAL, les éléments de recherche peuvent être pris dans différentes catégories d'informations mises en mémoire. Pour cette recherche les termes ont été recherchés parmi les descripteurs du "MESH annoté" qui recense le vocabulaire normalisé utilisable pour l'interrogation de MEDLARS - MEDLINE. Le MESH annoté comporte, comme le MESH ordinaire, des descripteurs et en plus des qualificatifs (subheadings) à affilier à ceux-ci pour en préciser l'emploi; son utilisation est particulièrement intéressante.

Dans l'interrogation du système MEDLINE on s'est servi de la pondération qui ne correspond pas à celle du système PASCAL VIRA. Dans le système MEDLINE le signe qui précède un descripteur pondéré met en évidence le caractère essentiel de la notion dans l'article correspondant. Ces descripteurs sont mis en relief sur le bordereau d'indexation; ils servent d'entrée dans l'Index Medicus et une recherche automatique qui porte exclusivement sur eux ne sélectionne que les articles concernant les notions qu'ils expriment. Leur utilisation diminue donc beaucoup le nombre de documents obtenus.

Les calculs de précision ont montré que MEDLARS a apporté davantage de bruit que PASCAL. En évaluant le rappel dans ces deux systèmes

pour cette même recherche on saura si le silence n'est pas supérieur dans le système PASCAL.

#### 74 Evaluation du rappel dans PASCAL et MEDLARS

Rappel =  $\frac{\text{Nombre de documents pertinents retrouvés}}{\text{Nombre de documents pertinents retrouvés} + \text{Nombre de documents pertinents non retrouvés.}}$

$$= \frac{A}{A + C}$$

La recherche a montré les résultats suivants:

Rappel dans le système PASCAL =	53,968 %
Rappel dans le système MEDLARS=	65,079 %

Le rappel est donc en effet meilleur dans le système MEDLARS que dans le système PASCAL.

#### 75 Conclusion

L'efficacité d'un système de recherche rétrospective dépend beaucoup de son degré d'évolution.

Dans le domaine de la biologie, le système PASCAL a pris dans sa progression un peu de retard sur les systèmes américains.

Les autres facteurs affectant les pourcentages de précision et de rappel d'un système de recherche rétrospective sont des facteurs intellectuels :

- Le langage d'indexation
- La politique d'indexation: le rappel augmente avec le nombre de descripteurs attribués à un document. Par contre si le nombre de descripteurs attribués à un document est faible le précision est souvent meilleure.

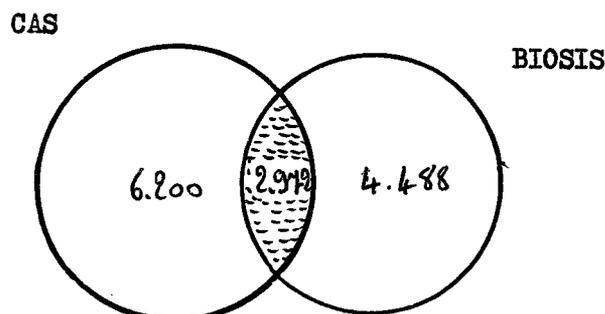
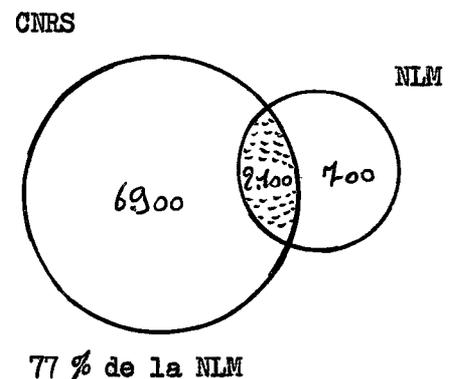
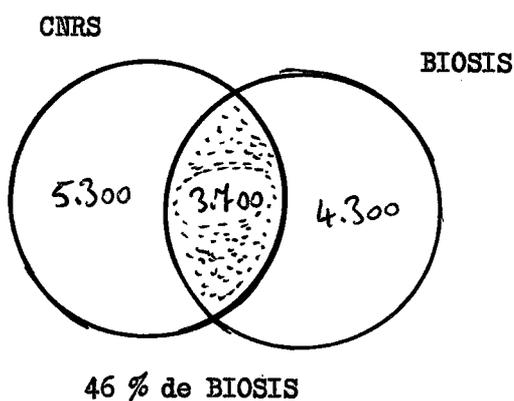
Puisque les anomalies de la précision sont généralement compensés par l'amélioration du rappel, le jugement <sup>porte</sup> sur un système dépend de

la priorité établie entre ces deux paramètres. Il est préférable d'obtenir davantage du bruit et moins de silence.

## 76 Couverture de la documentation par PASCAL, MEDLARS et BIOSIS

D'après une étude faite en 1975 sur le plan de la couverture scientifique, PASCAL couvre 2100 des 2800 revues traitées par MEDLARS, en outre il analyse des documents non périodiques. Mais comme PASCAL est pluridisciplinaire elle analyse des revues publiant des articles dans cette discipline sans y être spécifiquement consacrées. Au total PASCAL traite 3500 revues intéressant la médecine. Le secteur dentaire est mieux représenté dans MEDLARS que dans PASCAL.

Quant à situer PASCAL par rapport à BIOSIS, le CNRS traite environ 46 % des revues couvertes par BIOSIS. Donc en fonction de la loi de concentration de la littérature spécialisée en un nombre limité de revues, PASCAL signale la plus grande partie des travaux dans le domaine de la biologie. Mais BIOSIS accorde une large place à certains pays comme ceux de l'Amérique Latine.



- CAS couvre 40 % des titres de BIOSIS
- BIOSIS couvre 32 % des titres de CAS

## 8 COMPARAISON ENTRE PASCAL ET CA-COND

Introduction

Une tentative de comparaison entre PASCAL et CA-COND a été faite en 1978. Le travail a consisté à faire une recherche rétrospective en mode conversationnel dans les deux bases sur le sujet suivant: "Dosage par chromatographie phase gazeuse du dioxyde de carbone dans l'air".

La recherche a été faite sur deux serveurs: Sur Lockheed et SDC pour CA - COND et sur SDC uniquement pour PASCAL.

8I Evaluation du nombre de sorties, de la pertinence et du bruit

## Généralités

Pour évaluer la performance d'un système il est nécessaire d'étudier le nombre de références sorties, le taux de pertinence, le taux de bruit et le nombre de silences.

8II Etude du nombre de sorties

Les tableaux suivants montrent les résultats obtenus après interrogation et évaluation de la pertinence.

Tableau : Nombre de sorties lors de l'interrogation de CA-COND sur Lockheed et SDC

	ref. pertinentes	ref. marginales	ref. non pertinentes	total des ref.
Lockheed	6	4	11	21
SDC	7	1	2	10
Lockheed SDC	7	-	6	13
Total des ref. sorties	20	5	19	44

Type de ref.	ref. pertinentes	ref. marginales	ref. non pertinentes	total des ref.
Serveurs				
SDC	13	1	3	17

Tableau : nombre de sorties lors de l'interrogation de PASCAL sur SDC

	ref. pertinentes	ref. marginales	ref. non pertinentes	total des ref.
PASCAL	13	1	3	17
CA-COND	14	1	8	23

Tableau : Résultats de l'interrogation de PASCAL et de CA-COND sur SDC

En ce qui concerne la base CA-COND, le nombre de sorties, les taux de la pertinence et du bruit ont été calculés à partir des résultats obtenus lors de son interrogation sur le système SDC uniquement (tableau )

Il est clair que CA-COND a fourni un nombre de documents plus grand que celui de PASCAL, ceci paraît évident si on compare le nombre de revues sources spécialisées en chimie dépouillées par chacune.

## 812 Evaluation de la pertinence

Le taux de la pertinence de CA-COND a été de 45,4 % tandis que celui de PASCAL 76,4 %. Il est étonnant que le fichier CA-COND offre une pertinence moins élevée que celle de PASCAL, alors qu'à priori l'on pouvait s'attendre à l'inverse: le fichier CA-COND est en effet essentiellement orienté vers la chimie.

L'évaluation de la pertinence des documents a été pénible pour CA-COND parce que les signalements ne contiennent pas de résumés.

## 813 Calcul du bruit

De façon complémentaire au paragraphe précédent on a constaté ici que le taux de bruit dans PASCAL est de 17,6 % tandis que dans CA-COND il est de 34,4 %. Donc PASCAL a donné un taux de bruit beaucoup plus inférieur que celui de CA-COND.

Quant aux facteurs responsables du bruit ils pourraient être:

- Une mauvaise indexation .
- Un problème informatique au niveau de la connexion.
- Le bruit augmente lorsque l'équation de recherche est très simple et posée sans aucune restriction. Cependant , quelle que soit la qualité de la formulation , il est impossible d'éliminer totalement les bruits. Ici la question a été assez précise.

## 82 Langage d'interrogation

Tout comme le fichier PASCAL, CA-COND est interrogeable d'une part en langage naturel pour la zone titre et d'autre part en langage contrôlé pour la zone descripteurs.

## 83 Conclusion

Pour la question posée le fichier PASCAL a offert un meilleur taux de pertinence et de précision que le fichier CA-COND. Quant aux raisons de la médiocrité de ce dernier, elles sont difficiles à appréhender. Il semble que PASCAL est plus adaptée à cette question que CA-COND. Pour donner un

jugement objectif il serait souhaitable de faire l'étude à l'aide de plusieurs questions posées respectivement aux deux bases.

#### 84 Couverture de la documentation par PASCAL et CA-COND

Tout en essayant de couvrir le plus grand nombre de périodiques la base CA-COND se montre souvent sélective dans le dépouillement des articles des revues qu'elle traite; alors que PASCAL est beaucoup plus sélective dans le nombre de revues couvertes mais exhaustive dans leur dépouillement à cause de sa vocation pluridisciplinaire.

Dans les 2 systèmes les périodiques fournissent la majeure partie des références. Cependant dans CA-COND les brevets représentent 25 % du nombre total des références et proviennent de 26 nations alors que les brevets ne représentent que 2,6 % du total des signalements dans PASCAL et son pour la plupart d'origine française.

CAS est l'organisme le plus rapide dans le signalement des documents. Or le délai qui s'écoule entre la parution du document et son signalement dans les Chemical Abstracts est de 3 mois en moyenne alors que pour le Bulletin Signalétique il varie entre 4 à 6 mois.

Afin de pouvoir situer PASCAL par rapport à CA-COND et voir dans quelle mesure elle couvre la documentation en chimie il est indispensable de faire allusion au phénomène de concentration que Bradford a exprimé sous forme de loi: puisque l'unité documentaire dans le domaine des sciences et techniques est l'article de périodique, on constate qu'un nombre limité de revues de premier rang suffit pour couvrir la plus grande partie de la littérature parue dans un domaine donné. Donc pour avoir une idée sur la couverture de la documentation en chimie par PASCAL on a pointé sur la liste des 1000 premières revues le plus fréquemment citées par CA-COND : PASCAL dépouille 89 % des titres de cette liste.

En conclusion la base PASCAL assure une bonne couverture de la documentation en chimie et, lors de son interrogation en ligne,

elle a montré une bonne performance avec un taux de précision élevé. Les résultats obtenus par l'interrogation de CA-COND n'ont pas été très satisfaisants et il a été difficile de connaître les raisons.

## 9 EVALUATION DES DIFFERENTES BASES PAR L'AGENCE SPATIALE EUROPEENNE

Comme toutes les bases étudiées ici se trouvent sur l'ordinateur de l'ASE à Frascati, des études comparatives faites par l'ASE d'après le nombre de questions posées à chaque base, la pertinence des réponses et le degré de satisfaction des utilisateurs pourraient donner un jugement plus juste que des comparaisons à partir d'un nombre limité de questions choisies pour l'évaluation.

Ainsi le Centre de documentation de l'ASE a publié une liste de tous les domaines scientifiques couverts par les bases interrogeables à Frascati; cette liste s'appelle " Cosati Subject Catégories" (voir p. ) et est suivie d'un tableau indiquant une appréciation sur la couverture de chaque base dans chaque domaine cité dans la liste.

## CONCLUSION GENERALE

De cette analyse il ressort que plusieurs facteurs déterminent les résultats de recherche dans une base de données:

- . La connaissance des revues et des autres sources référées par les compilateurs de la base consultée et aussi la politique de sélection que ces compilateurs appliquent dans le choix des articles recensés.
- . La connaissance du langage d'indexation utilisé. Mais là il convient de noter qu'un fichier qui n'est interrogeable en langage libre que sur les mots du titre ne se comporte pas de la même façon qu'un fichier ayant, outre l'indexation libre une indexation contrôlée plus ou moins élaborée; d'où la nécessité de formuler différemment la question dans chaque fichier car aucun ne possède une structure identique à l'autre.
- . Il est à remarquer aussi que l'indexation par le vocabulaire libre uniquement a tendance à augmenter le silence. Ceci est dû au fait que souvent les mots du titre ne sont pas représentatifs du sujet traité dans l'article ; et que même s'ils le sont il est difficile pour le chercheur de penser à tous les synonymes qui expriment les différents aspects de son thème de recherche. L'usage d'un vocabulaire libre doit être un complément essentiel du vocabulaire contrôlé. De plus, tous les termes du vocabulaire d'indexation qu'ils soient contrôlés ou non doivent être enregistrés dans la mémoire de l'ordinateur afin de donner de bons résultats lors d'une recherche.

Jusqu'à présent une méthode idéale de recherche n'a pas été trouvée. Mais l'expérience a prouvé que si les systèmes sont bien utilisés et interrogés ils peuvent avoir de bonnes performances.

Le jugement de la pertinence des documents recensés lors de l'interrogation d'une base est souvent subjectif. Or dans certaines études comparatives on a constaté que le chercheur a rejeté certaines références alors qu'il en conservait d'autres assez similaires.

Dans le cas où le taux de bruit a été très important il a été difficile aux documentalistes d'en connaître la cause et seul un spé-

cialiste du sujet aurait pu les analyser . Cependant quelle que soit la qualité de la formulation de la question il est impossible d'éliminer totalement les bruits.

La recherche des silences montre que l'interrogation des fichiers varie en fonction de leur structure et que la pertinence seule ne permet pas de juger de l'efficacité d'un système. Si l'on désire être exhaustif il est préférable d'interroger au moins deux bases car la structure d'un fichier peut apporter l'apparition du silence, la nature et l'orientation de ce même fichier le peuvent également.

Quelle que soit la stratégie de recherche employée il est impossible de trouver tous les documents relatifs au thème de recherche et enregistrés dans la base . Ceci dépend , non seulement de la structure de la base mais aussi de la capacité du chercheur et du temps et de l'argent mis à sa disposition. Dans notre exemple de recherche dans PASCAL, INSPEC et COMPENDEX le pourcentage maximum des documents retrouvés était 72 %.

Le travail au terminal est cher et il est obligatoire de tenir compte du facteur coût pour l'interrogation. En temps d'ordinateur facturé on peut dire qu'un silence coûte plus cher qu'un bruit. Alors que ce dernier ne nécessite qu'une élimination manuelle, le premier entraîne des recherches supplémentaires au terminal.

La recherche avec une seule stratégie dans plusieurs bases est plus efficace que la recherche avec plusieurs stratégies dans une seule base.

Le plus grand nombre de documents recensés lors de la recherche bibliographique effectuée dans le cadre de notre étude traite de MEDLARS. Quant aux documents traitant des comparaisons entre les bases de données ils concernent presque tous des bases américaines uniquement sans que PASCAL y figure. Ceci pourrait être dû au fait que PASCAL n'est pas encore implantée sur le serveur du système Lockheed.

Le jugement de l'efficacité d'une base de données d'après l'interrogation sur un sujet choisi à priori est en effet hasardeux. Une base pourrait être très favorable pour une question et l'être beaucoup moins pour une autre dans la même spécialité. Par exemple dans l'étude comparative entre PASCAL et CA-COND, les résultats de CA-COND ont été médiocres alors que l'interrogation de cette même base a donné des réponses très satisfaisants lorsqu'elle a fait l'objet d'une comparaison avec d'autres bases américaines. De même en comparant les résultats des exemples de recherche présentés dans les pages précédentes avec le tableau de la couverture des bases (voir p. ) on découvre plusieurs contradictions par exemple:

D'après l'étude faite par l'ASE la couverture de CA-COND en agriculture est moyenne, alors que CA-COND a donné un bon résultat lorsqu'une question spécialisée en agriculture lui a été posée lors de sa comparaison avec BIOSIS, CAIN, AGRIS et SCI SEARCH.

L'efficacité des recherches dans les différentes bases dépend souvent du système d'interrogation (ou serveur). Autrement dit chaque serveur a un logiciel particulier qui offre des possibilités plus ou moins intéressantes selon la conception de chaque base.

L'étude précise des recouvrements de couverture de la documentation entre les différentes bases de données est si complexe qu'elle ne peut être réalisée que par une collaboration entre les producteurs des bases eux mêmes.

Les taux de recouvrement présentés dans notre travail montre que PASCAL couvrait en 1975 une bonne partie de la documentation déjà couverte par MEDLARS, INSPEC, BIOSIS et CA-COND;

- 77 % des revues déjà couverts par MEDLARS
- 62 % des revues déjà couverts par INSPEC
- 46 % des revues déjà couverts par BIOSIS
- 89 % des 1000 premières revues les plus citées par CA-COND.

Comme ces taux de recouvrement sont le résultat d'une étude effectuée en 1975 et approximative à 4 % il est nécessaire de reprendre actuellement cette étude afin de pouvoir suivre le développement de la base PASCAL dans ce sens.

Le taux de recouvrement des bases spécialisées est assez faible, ceci est évident vue la vocation de chacune.

A N N E X E S

cote CDST  
▷2

langues  
▷3

 P.A.S.C.A.L. I

▷V

ct.

ince

destinataires  
▷6

type  
▷7 **P**

type comp.  
▷8

niveau  
▷9

mode  
▷M

numéro bordereau

▷F

t. périodique  
▷1

ISSN  
▷B

pays  
▷O

date  
▷M

vol.  
▷V

no  
▷N

no d.  
▷H

pagination  
▷P

collation  
▷R

h.t.  
▷F

rés.  
▷K

bibl.  
▷L

ill.  
▷I

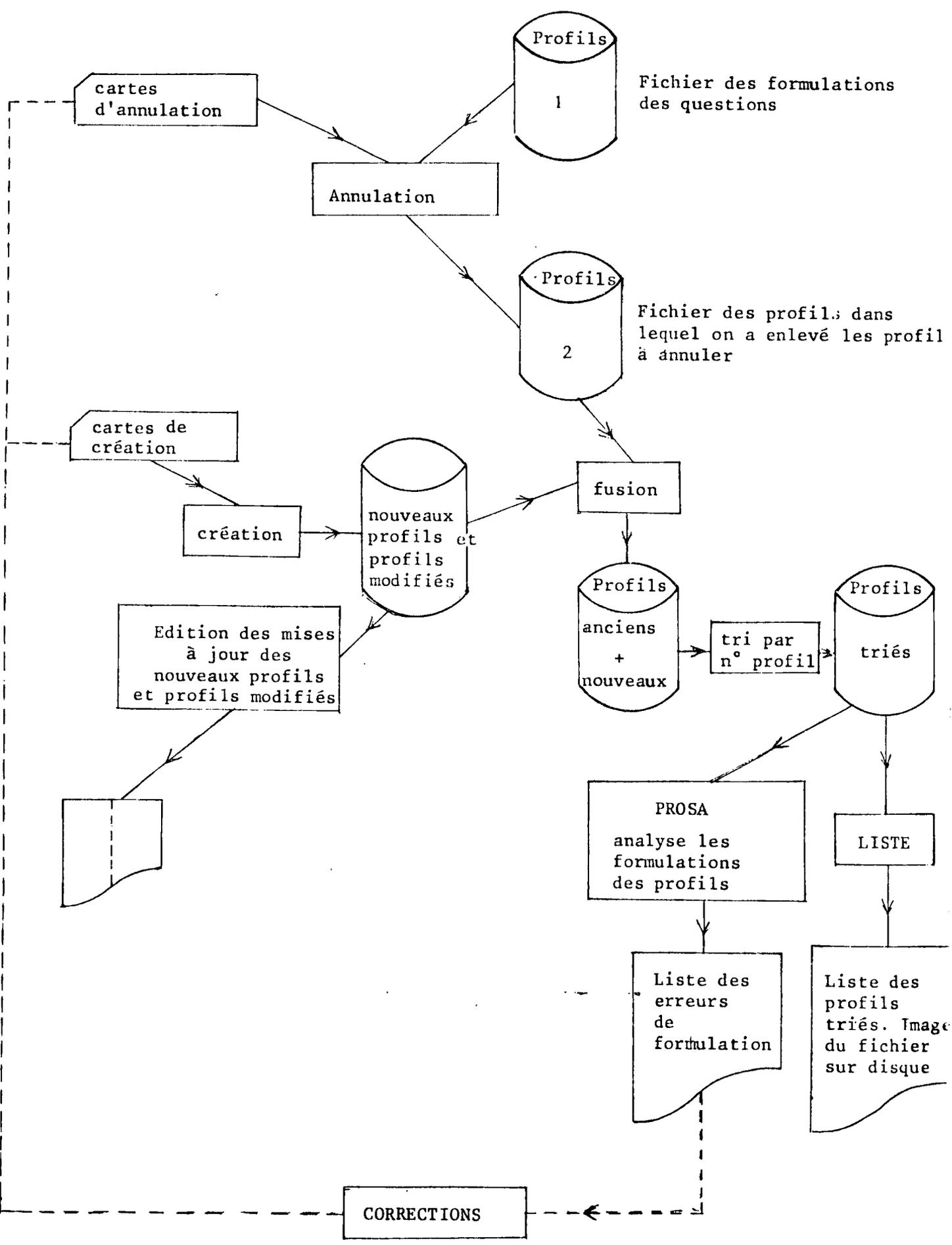
cart.  
▷C

anal.  
▷A

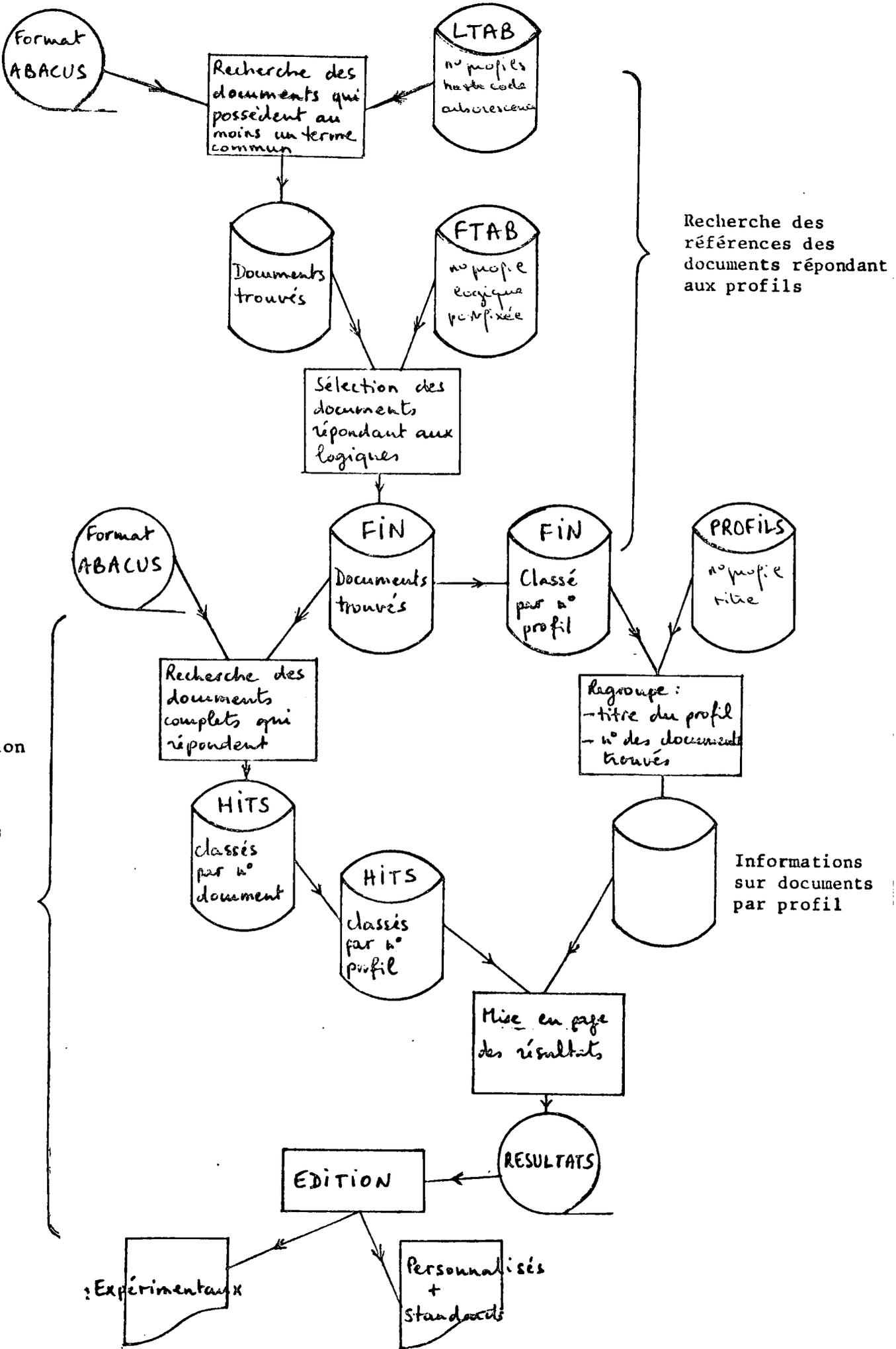
notes  
▷S

type comp.

▷F



DE  
EHE  
N TAIRE



opération  
dition  
s  
ultats

Recherche des  
références des  
documents répondant  
aux profils

Informations  
sur documents  
par profil



