

Thèse présentée à l'Université Lumière Lyon II pour l'obtention, le 02/10/1998, du  
**Doctorat de Sciences de l'Information et de la Communication**  
par Laurence FAVIER

**RECHERCHE ET APPLICATION D'UNE  
METHODOLOGIE D'ANALYSE DE  
L'INFORMATION POUR L'INTELLIGENCE  
ECONOMIQUE**  
**Application à un centre technique du secteur de la  
plasturgie**

Directeur de la thèse : Professeur Richard Bouché

Rapporteurs : Hubert Fondin, Ahmed Silem.

COMPOSITION DU JURY :

**Richard Bouché**, Professeur en sciences de l'information et de la communication, ENSSIB  
Villeurbanne

**Jean-Baptiste Carpentier**, Professeur en sciences de l'information et de la communication, Directeur  
du CELSA

**Jean-Pierre Fénelon**, Directeur de recherche au CNRS, statisticien.

**Hubert Fondin**, Professeur en sciences de l'information et de la communication, Université  
Bordeaux III

**Michel Kerbaol**, Ingénieur de recherche à l'INSERM, statisticien, Faculté de Médecine, Département  
de Santé Publique, Rennes

**Ahmed Silem**, Professeur en sciences de l'information et de la communication, Université de  
Lyon III.

**Jean-Pierre Soulier**, Professeur à l'Ecole Supérieure de Plasturgie, 01100 Bellignat.

Année universitaire 1997-1998

## Résumé français

Dans le cadre du développement d'un programme d'intelligence économique au sein d'un centre de recherche et développement d'une filière professionnelle (la filière plasturgie), cette thèse propose une méthodologie de traitement de l'information documentaire qui lui soit destinée.

Dans un premier temps (Partie I), une analyse de ce qu'est l'intelligence économique a été menée, afin d'élucider cette notion construite autour d'un anglicisme (le terme anglais *intelligence*). L'intelligence économique a pu être reliée à un héritage conceptuel et la manière dont elle a été importée tardivement en France a été mise en évidence.

Dans une seconde partie, cette étude présente les différentes méthodes de traitement de l'information documentaire utiles à l'intelligence économique. Elle propose ainsi une distinction entre recherche documentaire, bibliométrie et analyse de contenu; elle définit ensuite une méthode d'analyse de contenu appropriée à l'objectif recherché. Elle consiste dans l'application de l'Analyse Factorielle des Correspondances sur les textes des documents.

Après une comparaison entre cette méthode et celle que les américains appellent « Latent Semantic Analysis », la troisième partie expose une série d'expérimentations et aboutit à la proposition d'un système documentaire destiné à l'intelligence économique.

## Résumé anglais

This thesis takes place within the frame of a program on competitive intelligence in a research and development center of a professional branch (plastics processing). A methodology for information processing in this specific field is presented.

In this first part of the thesis, an analysis of what is competitive intelligence has been conducted in order to clarify the french denomination (built from the english term « intelligence »). The conceptual heritage of competitive intelligence has been outlined. It is shown how the concept has been (rather lately) introduced in France.

In the second part, a review of different methods of information processing relevant to competitive intelligence has been made. A distinction between information retrieval, bibliometrics and content analysis is proposed. A relevant method of content analysis is outlined. Correspondance Factor Analysis is applied on the texts of the documents. A comparison between this method and Latent Semantic Analysis has been conducted. The third part presents the experiments and leads to a proposal for a documentary system applied to competitive intelligence.

## **AVANT-PROPOS**

Cette thèse a été réalisée dans le cadre d'une convention CIFRE<sup>1</sup> de Novembre 1994 à Novembre 1997, liant le CERSI<sup>2</sup> et plus précisément l'axe "informatique documentaire" dirigé par le Professeur Richard Bouché et le Pôle Européen de Plasturgie. Elle a également fait l'objet d'une subvention d'EDF Entreprises Rhône-Alpes-Auvergne, grâce à Monsieur Bouland, responsable des grands comptes.

Je tiens à remercier ceux qui ont permis à cette thèse de voir le jour et en ont garanti l'accomplissement : Monsieur Bouland qui a contribué à son financement avec l'ANRT, le Professeur R.Bouché qui a veillé à la réussite de la convention, le Directeur Général du Pôle Européen de Plasturgie (de sa création à Juin 1998) et l'ancien président du Pôle (Monsieur Paturle) jusqu'en 1997 qui a initié cette recherche et en a souhaité la continuation.

Je voudrais également témoigner toute ma reconnaissance à Monsieur Michel Kerbaol, statisticien au Département Santé Publique de la Faculté de Médecine de Rennes, qui a bien voulu mettre à ma disposition ses moyens de calculs et contribué à la mise en oeuvre de mes travaux ainsi qu'à la réflexion qui les a animés. Je remercie également son collaborateur, Monsieur Jean-Yves Bansard, pour avoir participé à cette réflexion.

---

<sup>1</sup> Convention Industrielle de Formation par la Recherche : ce type de convention lie un laboratoire et une entreprise, ici une association d'entreprises, et bénéficie d'une subvention de l'ANRT (Association Nationale pour la Recherche et la Technologie) pour trois années.

<sup>2</sup> Centre de Recherche et d'Etudes en Sciences de l'Information, laboratoire de l'Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques jusqu'en Octobre 1998.

# SOMMAIRE

<b>INTRODUCTION GENERALE .....</b>	<b>8</b>
<b>A- EMERGENCE DE L'INTELLIGENCE ECONOMIQUE .....</b>	<b>9</b>
<b>B- LA DEMANDE DU PEP .....</b>	<b>12</b>
a) LE CONTEXTE : LE POLE EUROPEEN DE PLASTURGIE (PEP).....	12
b) L'INTELLIGENCE ECONOMIQUE .....	14
c) PRESENTATION DU PLAN DE CETTE ETUDE.....	15

## **PARTIE I : VEILLE ET INTELLIGENCE ECONOMIQUE : ELUCIDATION THEORIQUE DES CONCEPTS**

<b>INTRODUCTION.....</b>	<b>17</b>
<b>I-1 ELUCIDATION DES TERMES.....</b>	<b>18</b>
I-1-1 ANALYSE DES OCCURRENCES .....	18
I-1-2 APPARITION AUX USA DU THEME DE LA SURVEILLANCE DE L'ENVIRONNEMENT .....	27
Les premiers documents sur "environmental scanning system"/système de surveillance de l'environnement .....	27
I-1-3 APPARITION DU THEME INTELLIGENCE ECONOMIQUE .....	32
Les premiers documents sur "competitive/business intelligence" .....	32
I-1-3.1 Information scientifique et technique et intelligence économique.....	32
I-1-3.2 Intelligence économique et management stratégique .....	33
I-1-3.3 Intelligence économique et organisation .....	35
I-1-4 APPARITION DE CES NOTIONS EN FRANCE .....	40
Premières références françaises.....	40
I-1-4.1 La veille.....	40
I-1-4.2 L'intelligence économique.....	41
I-1-5 RESULTATS DE CETTE COMPARAISON DE L'USAGE DES TERMES EN FRANCE ET AUX USA.....	42
<b>I-2 ELUCIDATION DES CONCEPTS .....</b>	<b>44</b>
I-2-1 SURVEILLER L'ENVIRONNEMENT : LES MODELES THEORIQUES UTILISES .....	44
I-2-1.1 Le management stratégique et le thème de l'analyse de l'environnement .....	46
L'Analyse par l'activité .....	46
Analyse par les forces concurrentielles .....	51
Analyse par les processus de décision.....	56
Conclusion.....	59

I-2-1.2 La contestation de la planification stratégique : discussion sur le rôle et la nature de l'information dans la décision.....	60
a) Le point de vue de la gestion : "Les hypothèses de base cachées derrière la planification stratégique" .....	60
b) L'apport de la sociologie .....	62
I-2-2 LA GESTION DE LA TECHNOLOGIE.....	63
CONCLUSION.....	65
<b>I-3 ROLE DE LA "FONCTION DOCUMENTATION" ET DES SYSTEMES DOCUMENTAIRES ..</b>	<b>65</b>
I-3-1 DOCUMENTATION ET INTELLIGENCE ECONOMIQUE.....	65
I-3-2 ETUDE DE 1995 SUR LE THEME "COMPETITIVE INTELLIGENCE".....	66
I-3-3 CONSTAT SUR L'ORGANISATION DES STRUCTURES DE DOCUMENTATION EN France...	68
I-3-4 EVOLUTION DE LA DOCUMENTATION .....	68
<b>I-4 CONCLUSION.....</b>	<b>69</b>
<b>ANNEXE I-A .....</b>	<b>71</b>
<b>ANNEXE I-B .....</b>	<b>72</b>
<b>PARTIE II- LE TRAITEMENT DE L'INFORMATION TEXTUELLE, OUTIL DE L'INTELLIGENCE ECONOMIQUE .....</b>	<b>73</b>
<b>II-1 LE PROBLEME .....</b>	<b>74</b>
II-1 -1 QU'EST-CE QU'UN DOCUMENT ?.....	74
II-1-2 LA VALEUR STRATEGIQUE DES DOCUMENTS.....	76
<b>II-2 LES TRAITEMENTS POSSIBLES DU DOCUMENT TEXTUEL.....</b>	<b>77</b>
II-2-1 LA RECHERCHE DOCUMENTAIRE.....	78
II-2-2 L'ANALYSE DE CONTENU.....	78
II-2-3 LA BIBLIOMETRIE .....	79
<b>II-3 L'INFORMATIQUE DOCUMENTAIRE : .....</b>	<b>80</b>
II-3-1 LES METHODES.....	800
II-3-2 LES LIMITES .....	82
<b>II-4 STATISTIQUE SUR LES DONNEES TEXTUELLES (ANALYSE DE CONTENU) OU SUR LES DONNEES SECONDAIRES DU DOCUMENT (BIBLIOMETRIE) : .....</b>	<b>85</b>
II-4 -1 LA BIBLIOMETRIE, PRESENTATION GENERALE.....	85
II-4-2 DE LA BIBLIOMETRIE DISTRIBUTIONNELLE A LA BIBLIOMETRIE DES REFERENCES BIBLIOGRAPHIQUES .....	877
II-4-2. 1 Les lois bibliométriques : .....	877
II-4-2. 2 - La création de l'ISI : de la gestion de la documentation à l'évaluation de la recherche .....	900
II-4-2. 3 Les indicateurs .....	911
II-4-2. 4 L'analyse des associations .....	933
II-4-2. 4,1 : Les objets .....	93
II-4-2. 4, 2 Les méthodes.....	944
a) les classifications.....	95
b) L'analyse factorielle.....	100

II-4-3 DE LA SCIENTOMETRIE A L'ANALYSE DE CONTENU .....	1000
II-4-3. 1 Les classifications hiérarchiques .....	1011
II- 4- 3. 2 L'analyse factorielle .....	101
II- 4- 3. 1, 1 Analyse Factorielle des Correspondances : .....	101
a) Présentation de ces trois points : .....	102
b) L'hypothèse linguistique de l'AFC .....	103
II- 4- 3 1, 2 Latent semantic analysis (LSA) ; Latent Semantic Indexing (LSI) .....	109
a) introduction .....	109
b) La méthode de LSA : Singular Value Decomposition : SVD .....	110
c) Présentation de Latent Semantic Analysis .....	112
d) Comparaison entre AFC ET LSA .....	115

**PARTIE III. PROPOSITION POUR UNE METHODE D'ANALYSE DE L'INFORMATION TEXTUELLE DESTINEE A L'INTELLIGENCE ECONOMIQUE : EXPERIMENTATIONS ..... 127**

**III-1 MATERIEL ET METHODE ..... 128**

III-1-1 LE CHOIX DE L'ANALYSE FACTORIELLE DES CORRESPONDANCES (AFC) POUR TRAITER LE TEXTE LIBRE .....	129
III-1-1. 2 La démarche .....	129
III-1-1. 3 Le problème méthodologique .....	131
III-1-1. 4 Les motivations de l'utilisation de la méthode.....	132

III-1-2 LES TEXTES TRAITES.....	133
---------------------------------	-----

III-1-3 METHODE .....	133
-----------------------	-----

III-1-4 LOGICIEL UTILISE.....	134
-------------------------------	-----

III-1-5 PRESENTATION DES RESULTATS .....	135
--	-----

**III-2 RESULTATS ..... 137**

INTRODUCTION .....	137
--------------------	-----

III-2-1 LA COHERENCE DES ASSOCIATIONS .....	137
---	-----

III-2-1.1 Interprétation par "métaclé" .....	137
--	-----

<sup>2</sup> III-2-1.2 Des métaclés à la typologie thématique.....	147
--	-----

III-2-1.3 Les déceptions de l'interprétation.....	147
---	-----

III-2-1.4 Les cartographies thématiques.....	155
--	-----

III-2-2 INFLUENCE DU PRINCIPE DE SELECTION DES UNITES STATISTIQUES SUR LES RESULTATS .....	157
--	-----

III-2-2.1 Comparaison des trois analyses entre elles .....	159
--	-----

III-2-2.2 Comparaison des typologies thématiques obtenues avec l'analyse 1 (746) et l'analyse 2 (673).....	161
--	-----

III-2-2.3 Les différences avec l'analyse 296 .....	165
--	-----

III-2-2.4 Conclusion .....	165
----------------------------	-----

III-2-3 ANALYSE DE DOCUMENTS REJETES .....	166
--	-----

III-2-4 LES NOTIONS DE SYNTHESE ET DE THEME.....	167
--	-----

Comparaison entre une synthèse issue d'un groupe d'experts et une synthèse obtenue par AFC	
--	--

sur les mêmes documents

III-2-5 CIBLAGE D'UNE EVOLUTION CHRONOLOGIQUE .....	169
<b>III-3 DISCUSSION DES RESULTATS .....</b>	<b>172</b>
III-3-1 LA DÉSAMBIGUÏSATION SANS DICTIONNAIRE.....	172
III-3-2 LA LEMMATISATION RENDUE CADUQUE .....	172
<b>III-4 DIFFUSION DES RESULTATS. Quel système documentaire pour l'intelligence économique ?</b>	<b>178</b>
<b>III-5 CONCLUSION.....</b>	<b>180</b>
<b>CONCLUSION GENERALE .....</b>	<b>181</b>
<b>RESULTATS DETAILLES .....</b>	<b>187</b>
<b>ANNEXE III-A .....</b>	<b>221</b>
<b>ANNEXE III-B .....</b>	<b>222</b>
<b>ANNEXE III-C .....</b>	<b>223</b>
<b>BIBLIOGRAPHIE .....</b>	<b>227</b>
<b>LISTE DES TABLEAUX ET DES FIGURES .....</b>	<b>237</b>

# **INTRODUCTION GENERALE**

# RECHERCHE ET APPLICATION D'UNE METHODOLOGIE D'ANALYSE DE L'INFORMATION POUR L'INTELLIGENCE ECONOMIQUE.

Application dans un centre technique du secteur de la plasturgie

## INTRODUCTION GENERALE

### ***A- EMERGENCE DE L'INTELLIGENCE ECONOMIQUE***

On pourrait définir l'intelligence économique comme le renseignement au service des acteurs économiques : elle consiste dans la mise en place d'un système permettant d'obtenir des renseignements sur ses concurrents, plus généralement sur son environnement concurrentiel (évolutions technologiques, scientifiques, du marché, et connaissance des acteurs qui les animent). On l'oppose à l'espionnage économique, l'intelligence économique étant le traitement légal de l'information alors que l'espionnage en est l'acquisition illégale. Cette notion "d'intelligence économique", bien qu'ancienne comme nous le verrons au cours de la première partie de notre étude, est aujourd'hui d'une actualité telle qu'elle préoccupe aussi bien les entreprises, les organisations professionnelles que les Etats et instances supranationales.

Depuis le début des années 1990 la question de l'intelligence économique connaît un succès considérable en France : articles de presse, ouvrages, rapports officiels, formations à l'intelligence économique se multiplient, sans oublier les conséquences de ce succès sur les métiers impliqués à divers titre dans cette question (on voit apparaître des prestations en intelligence économique de la part de ceux qui vendaient des services de veille technologique ou de documentation).

En 1994, quelques mois avant que ne débute cette étude, le rapport du Commissariat Général au Plan (XI<sup>ème</sup> Plan), consacrait le terme "d'intelligence économique" en intitulant son rapport "Intelligence économique et stratégie des entreprises". Ce rapport se situait dans la continuité des réflexions menées par le Commissariat Général du Plan au cours des années passées sur les facteurs immatériels de la compétitivité. Cet anglicisme, traduction du terme "competitive intelligence", marque le terme d'une évolution qui, en France, a débuté avec l'**information scientifique et technique** et les politiques qui lui ont été consacrées. On a ensuite

"institutionnalisé" le terme de "**veille technologique**" avec en 1988 la création du comité "d'Orientation Stratégique de l'Information scientifique et de la Veille technologique" par le Ministre de la Recherche Jacques Valade, puis avec le groupe "Veille technologique et politique des brevets" de la sous-commission "Innovation et Recherche" en 1989. Enfin les travaux du groupe "Information et compétitivité" puis ceux de la commission "Compétitivité française" du XI<sup>ème</sup> Plan ont abouti à "**l'intelligence économique**".

Notre travail s'est donc inscrit dans ce nouveau projet : l'intelligence économique et la stratégie des entreprises. Elle se définissait en 1994 comme une notion qui implique : "*le dépassement des actions partielles désignées par les vocables de documentation, de veille (scientifique, technologique, concurrentielle, financière, juridique et réglementaire...), de protection du patrimoine concurrentiel, d'influence (stratégie d'influence des Etats-Nations, rôle des cabinets de consultants étrangers, opérations d'information et désinformation...).*" [Plan, 1994, p. 16-17]

Le besoin de coordonner des "*actions partielles*", ou plus exactement la nécessité ressentie de lier dans un même objectif ce qui, jusqu'à présent, avait été pensé de façon isolée justifie l'appel à la notion "*d'intelligence économique*".

Le Rapport poursuit : "*Ce dépassement résulte de l'intention stratégique et tactique, qui doit présider au pilotage des actions partielles et au succès des actions concernées, ainsi que de l'interaction entre tous les niveaux de l'activité, auxquels s'exerce la fonction d'intelligence économique : depuis la base (internes à l'entreprise) en passant par des niveaux intermédiaires (interprofessionnels, locaux) jusqu'aux niveaux nationaux (stratégies concertées entre les différents centres de décision), transnationaux (groupes multinationaux) ou internationaux (stratégie d'influence des Etats-Nations).*" [Plan, 1994, *ibid.*]

La coordination des actions partielles de documentation, veille, protection, influence, etc. au service d'une "*intention stratégique et tactique*" signifie donc que le problème n'est plus la satisfaction des besoins d'information des agents économiques ou des Etats selon des objectifs indépendants mais que "*la gestion stratégique de l'information économique*" en tant que "*l'un des moteurs essentiels de la performance globale des entreprises et des nations*" [*ibid.*, Avant-Propos] repose sur "l'interaction" des niveaux de l'entreprise, de la profession, de la nation, des instances supranationales.

Le "Livre vert sur l'innovation" de l'Union Européenne reprend à son compte en 1996 le problème de l'intelligence économique ainsi défini [Livre Vert, 1996, p. 21]. Il la présente comme un "*outil stratégique d'aide à la décision dans un contexte de mondialisation des échanges et d'émergence de la société de l'information*" (*ibid.*). Mais, outre ce contexte, c'est le constat suivant qui motive l'appel à "l'intelligence économique" : "***Paradoxalement, l'offre croissante de données, grâce aux technologies de l'information ne se traduit pas en une perception supérieure des enjeux technologiques et économiques ni en une plus grande visibilité des choix stratégiques à opérer***".

L'intelligence économique est donc ce qui doit permettre de résoudre un tel paradoxe. Son action se situe à un double niveau : celui de la firme et des organisations professionnelles d'une part, celui de la politique nationale et supranationale en matière d'information d'autre part. La veille technologique et concurrentielle, préconisant une surveillance de l'environnement, semblait se limiter au seul niveau de l'entreprise (l'entreprise confrontée à son environnement) face à laquelle l'Etat devait s'employer à développer l'offre d'information scientifique et technique et la politique de formation aux Technologies de l'Information et de la Communication (dites TIC). Le résultat d'une telle approche serait donc encore insuffisant puisque "l'offre croissante de données" n'améliore ni la "perception" des enjeux ni la "visibilité" des choix stratégiques.

Nous nous sommes alors proposés d'analyser ce problème :

- d'une part, en cherchant à comprendre sa nature et son évolution ainsi que le rôle que pouvaient jouer les systèmes documentaires dans ce contexte.

- d'autre part, en répondant à une demande du Pôle Européen de Plasturgie, qui souhaitait en 1994, que lui soit proposée une approche permettant d'effectuer une analyse de l'information documentaire résolvant la difficulté résumée dans le paradoxe ci-dessus décrit par le Livre Vert : à savoir exploiter "l'offre croissante de données" par une méthode d'analyse de l'information susceptible d'améliorer "la perception des enjeux" et la "visibilité des choix". **En d'autres termes, comment améliorer les méthodes de recherche et de traitement de l'information documentaire pour répondre, non plus à la recherche de données, mais à des questions stratégiques ?**

## **B- LA DEMANDE DU PEP**

### **a) LE CONTEXTE : LE POLE EUROPEEN DE PLASTURGIE (PEP)**

Il s'agit d'une association de 170 entreprises impliquées dans la plasturgie c'est-à-dire la transformation thermo-mécanique des polymères.

Elles représentent :

- plusieurs secteurs : les fournisseurs (chimistes pétroliers), les intégrés (notamment les principaux donneurs d'ordre : automobile, aérospatial), les transformateurs (fabricants de pièces plastiques), les outilleurs (moulistes et secteur de la mécanique), les fabricants de machines, l'informatique (logiciels de simulation des procédés, de CAO et de rhéologie), l'électronique (mesures physiques), le design (pour la conception de l'objet), les traitements de surface des polymères, les banques.
- plusieurs types d'entreprises, différentes par leur taille et modes d'organisation, même si l'essentiel de la plasturgie est constitué de petites et moyennes entreprises.

Le P.E.P se distingue d'un centre technique tout d'abord par son mode de financement qui est volontaire. Il ne repose pas sur un système du type "taxe parafiscale" mais sur les cotisations des entreprises d'une part et sur les contrats d'études d'autre part. Par ailleurs, ce centre est l'une des trois entités d'un pôle comprenant en outre une école d'ingénieurs (école publique) et un futur Centre d'Entreprises et d'Innovation. Ce type de structure constitue une différence supplémentaire avec les centres techniques.

Le centre de documentation n'a pas seulement pour vocation de mettre à disposition des ressources documentaires pour ses adhérents et pour la formation, mais doit pouvoir également dispenser des prestations à ses adhérents.

De plus, il a pour vocation de répondre à tous les publics de cette structure et donc à des besoins d'information différents.

Les missions du PEP sont les suivantes :

- Recherche et développement
- Essais
- Formation (initiale et continue)
- Documentation

L'activité "intelligence économique" a été liée à la documentation.

- La documentation

L'une des missions de cette structure est de constituer un service de documentation pour la profession de la plasturgie, ce qui n'existait pas jusqu'à présent, la documentation de ce domaine étant éparpillée dans les différents secteurs industriels qui utilisaient notamment la plasturgie. En effet la recherche d'informations dans ce domaine présente des difficultés particulières liées à :

- un manque de "lisibilité" du mot de *plasturgie* aussi bien en français qu'en anglais. Peu de bases de données sont spécialisées dans ce domaine.
- la pluridisciplinarité, la filière plasturgie étant au cœur des interactions entre matériaux polymères, outillages, technologies et procédés, captage et traitement de l'information *in situ*, design, conception de pièces.

Une fois prise en compte le besoin de collecter de l'information pour la plasturgie par la création d'un centre de ressources restait à répondre à un autre besoin : celui d'études stratégiques sur le domaine. Notre mission a été de proposer des outils pour répondre à ce besoin.

La profession, spécifiquement celle des transformateurs évolue vers une prise en charge de sa propre recherche et développement c'est-à-dire une indépendance vis à vis de leurs fournisseurs et leurs donneurs d'ordre, ce qui amène la profession à de nouveaux besoins en matière d'information. Cette évolution très importante conduit au développement de l'intelligence économique autant qu'à celui de la recherche et du niveau de formation des cadres ; elle est au fondement même de la création, souhaitée par les industriels, du Pôle Européen de Plasturgie (PEP).

## b) L'INTELLIGENCE ECONOMIQUE

Dans le cadre de la mise en place d'une activité d'intelligence économique au PEP, notre tâche a été de proposer une approche permettant :

- le traitement d'une grande quantité d'informations pour répondre à des besoins d'informations stratégiques (ciblage des thèmes principaux constituant un sujet, de concepts nouveaux, évolution au cours du temps d'un domaine, etc.).
- un système de diffusion et de consultation de l'information primaire et de l'information traitée adapté à la fois aux besoins internes de la structure et aux besoins des membres externes.

La collecte des données, la connaissance des sources d'informations fait partie de l'intelligence économique autant d'ailleurs que les questions liées à la sécurité. Mais nous ne les avons pas abordées dans cette étude. Cette connaissance des sources d'information externe repose sur la compétence des documentalistes. Les questions relatives à la sécurité de l'information sont souvent l'affaire de spécialistes. Nous avons limité notre apport en la matière à l'un des aspects essentiels de l'intelligence économique : celui du traitement des documents.

Le traitement de l'information documentaire est classiquement celui de la synthèse élaborée éventuellement avec un groupe d'experts à partir des sources consultées de façon empirique car rien ne permet de trier a priori de façon objective la surabondance d'information. Ce type de traitement reste assez irremplaçable en matière d'état de l'art afin de prendre connaissance d'un sujet dont on n'est pas spécialiste. Mais il est notoirement insuffisant pour les études stratégiques. C'est la raison pour laquelle une méthodologie et des outils spécifiques doivent être réalisés à cette fin.

Le système de diffusion de l'information documentaire échappe très souvent aux documentalistes pour être pris en charge par les services informatiques. Il est en effet de plus

en plus indispensable que le système documentaire soit inséré dans le système d'information général de l'entreprise (lequel relève de la responsabilité des informaticiens de la structure), sans être relégué à une gestion électronique des documents que seul le documentaliste pourrait manipuler. La gestion électronique des documents, surtout à l'heure des réseaux, est d'ailleurs conçue pour répondre à un besoin beaucoup plus global que celui des documentalistes : il répond à la gestion de tout type de document dans l'entreprise. Il reste que la circulation non plus des documents *stricto sensu* mais des connaissances véhiculées par ceux-ci (des produits d'information plus ou moins élaborés répondant à un besoin précis) doit être pensée en tant que telle. Cette tâche relève également de l'intelligence économique.

#### c) PRESENTATION DU PLAN DE CETTE ETUDE :

**La première partie** de notre travail est consacrée à l'élucidation des notions de veille technologique et d'intelligence économique.

L'objectif de cette première partie est de procéder à une analyse de l'intelligence économique qui ne nous semble pas avoir été effectuée jusqu'à présent et de comprendre pourquoi des méthodes documentaires peuvent légitimement répondre à des enjeux stratégiques.

**La deuxième partie :** passera en revue les différentes approches en matière de traitement de l'information documentaire textuelle.

**La troisième partie** proposera une méthodologie pour un traitement de l'information adapté à l'intelligence économique à travers des exemples d'études réalisées dans le cadre de la mission au P.E.P et s'achèvera sur une proposition de système documentaire permettant d'exploiter cette méthodologie.

**Partie I**  
**VEILLE ET INTELLIGENCE ECONOMIQUE :**  
**ELUCIDATION THEORIQUE DES CONCEPTS**

# PARTIE I : VEILLE ET INTELLIGENCE ECONOMIQUE : ELUCIDATION THEORIQUE DES CONCEPTS

## **INTRODUCTION**

La définition des concepts de veille et d'intelligence économique a été soigneusement fixée par les rapports officiels, de multiples écrits ont été consacrés à ces sujets. Nous aurions pu alors considérer que ces concepts étaient suffisamment clairs pour ne pas avoir à y revenir, ou tout au plus, pour n'y consacrer que de brefs rappels. Mais une grande confusion nous a paru entourer ces concepts. Bien que s'inspirant de termes et de modèles américains clairement identifiés, leur traduction semble être difficile. Ainsi, si nous traduisons par "veille technologique" le terme "environmental scanning", nous serons mal compris des américains qui lui préféreront le terme "competitive intelligence". Or les Français traduisent cette dernière expression par "intelligence économique". Nous proposant précisément d'examiner un certain nombre de ces difficultés terminologiques liées au passage d'une langue à l'autre, nous ne traduirons donc pas, dans les démonstrations qui suivront, les termes anglo-saxons.

Au-delà de ces difficultés terminologiques, nous sommes confrontés en France à deux discours sur ces thèmes : celui des sciences de gestion et celui des sciences de l'information, et plus précisément de la documentation. Les prestataires de service en veille ou en intelligence économique proposent aux industriels tantôt des méthodes d'analyse du positionnement concurrentiel baptisées "veille technologique et concurrentielle"<sup>3</sup> tantôt des méthodes documentaires de diffusion sélective d'information nommée soit "veille technologique", soit "veille stratégique", ou encore "veille concurrentielle" selon le domaine de l'information recherchée. La confusion demeure pour l'industriel intéressé par la "veille" quand il se trouve confronté aussi bien aux ouvrages de Lesca qu'à ceux de Jakobiak, tous deux pionniers en la matière en France. Qu'y a-t-il de commun entre le "Système d'information pour le management stratégique (SIMSE)" de l'un [Lesca, 1986], et la "Structure de Surveillance Sectorielle Systématique" de l'autre [Jakobiak, 1991] ? Quel modèle partagé par ces deux approches, l'une se réclamant de la gestion stratégique, l'autre de l'informatique documentaire ? Si l'organisation d'un service documentaire et ses méthodes d'analyse

---

<sup>3</sup> Voir par exemple la méthode INITEC développée par l'Ecole Supérieure de Commerce de Grenoble, et proposée en stage de formation continue par l'ARIST Rhône-Alpes (stage du 23-26/10/1995, "Le projet de veille technologique et concurrentielle").

participent à la veille sans s'y réduire, qu'est-ce que la veille précisément ? S'il suffit de penser en termes de management stratégique pour faire de la veille, en quoi a-t-on besoin de la notion de "système d'information" ?

Nous avons cherché à élucider les notions de veille et d'intelligence économique sans nous placer d'emblée dans l'un ou l'autre champ disciplinaire (sciences de gestion ou documentation). Notre objectif est d'abord de comprendre quels problèmes sont liés à ces notions.

La réponse en termes de méthode documentaire que nous proposerons en deuxième partie se situera dans le prolongement de cette réflexion générale, et non comme une proposition répondant à un choix d'hypothèse sur la nature de la veille ou de l'intelligence économique (l'alternative sciences de gestion/science de la documentation étant elle-même susceptible d'être interrogée).

La double élucidation terminologique et conceptuelle à laquelle nous nous sommes livrés nous a paru d'autant plus indispensable que nous avons assisté durant ces trois années d'étude à une explosion de l'intérêt porté à l'intelligence économique : formations, ouvrages, articles de presse... Or nous nous sommes aperçus, avec un certain étonnement, que le problème avait pourtant une trentaine d'années. Pour quelle raison voit-on alors se développer un intérêt récent pour un problème qui ne l'est pas ? Les réponses et les justifications ne manquent pas mais elles ne résistent pas à l'examen, nous le montrerons dans la partie I-1-1 concernant le commentaire des résultats.

## ***I-1 ELUCIDATION DES TERMES***

### **I-1-1 ANALYSE DES OCCURRENCES**

Les termes de "veille" et "d'intelligence économique" sont des reprises explicites de termes américains. Le premier est la traduction de "environmental scanning system" (parfois également "strategic information scanning system" [Aaker, 1983]) ; le second constitue un anglicisme autour du terme (devenu par-là même ambigu) "d'intelligence". Rappelons que "intelligence" en français n'a jamais la signification de "renseignement" mais renvoie à l'idée

de faculté de comprendre ; d'où une certaine confusion liée à cette expression "d'intelligence économique" en français.

Au-delà des termes, que pouvons-nous dire quant à la formulation des problèmes qui leur sont attachés de part et d'autre de l'Atlantique ?

Un premier constat s'impose : l'évolution du vocabulaire qui nous fait passer progressivement, en France, de la veille technologique à l'intelligence économique **ne se retrouve pas** dans la littérature américaine.

Nous l'avons constaté à partir de l'interrogation du Dialindex de DIALOG pour vérifier l'utilisation des termes dans les bases américaines. Le Dialindex permet d'interroger un ensemble de bases simultanément, ensemble que l'on peut choisir dans un certain nombre de catégories préconstituées par thème. Les bases ayant le plus grand nombre d'occurrences des termes nous intéressant ont été sélectionnées à l'intérieur d'une catégorie rassemblant un ensemble de bases concernant l'industrie (MULTIIND). Nous avons également interrogé deux bases de données françaises, DELPHES et FRANCIS, afin de connaître les premiers emplois des termes de "veille" et "d'intelligence économique" et comparer ainsi les résultats avec ceux issus de l'interrogation des bases de données américaines.

Nous avons enfin consulté trois bases en sciences sociales : SOCIAL SCISEARCH, SOCIOLOGICAL ABSTRACTS et ECONOMIC LITERATURE INDEX.

L'objectif de notre démarche n'est pas de rechercher de façon exhaustive toutes les références sur le sujet mais de détecter quelques tendances dans l'emploi du vocabulaire lié au problème qui nous préoccupe.

### **Les bases consultées :**

Au total, ce sont neuf bases de données qui ont été consultées. Nous en donnons une description rapide dans les lignes suivantes.

#### ABI/INFORM :

Domaine : économie et gestion des entreprises

Données : articles de 1200 périodiques, à 87% américain dont : Harvard Business Review, Dun Business Month, Fortune, Forbes, Journal of Marketing Research, Management today (Grande Bretagne), International Labour Review (Suisse).

### IAC PROMT :

Domaine : activité des entreprises, industries, commerce, marchés de toutes entreprises, tous secteurs confondus.

Données : 1200 journaux, revues et études du monde entier : Wall Street Journal, Chemical Marketing reporter, Forbes, Electronic news, Chemical and Engineering news, Research and Development, Corporate industry government and agencies research reports, Investex reports, rapports annuels de sociétés.

### IAC TRADE & INDUSTRY DATABASE :

Domaine : Affaires, commerce et industries, Etats Unis.

Données : 960 périodiques dont Fortune et Money area business databank, et les dépêches de l'agence de presse PR Newswire

### TEXTLINE GLOBAL NEWS :

Domaine : actualité, économie, entreprises, biographies. International, mais l'accent est mis sur l'Europe.

Données : issues des principales publications économiques européennes et des agences de presse (dont Reuter News Service).

### SOCIAL SCISEARCH :

Domaine : sciences humaines et sociales, médecine

Données : références bibliographiques issues des articles de 1500 périodiques et, par ailleurs, 3000 périodiques de sciences exactes et médicales (références intéressant les sciences sociales).

### SOCIOLOGICAL ABSTRACTS :

Domaine : sociologie et discipline voisines

Données : 1900 périodiques, comprend également les communications à des congrès. Couverture mondiale, dont 60% USA.

### ECONOMIC LITERATURE INDEX :

Domaine : économie, financement et investissement, marketing, organisation industrielle, changement technologique, consommation.

Données : issues de 300 périodiques importants du monde économique.

### DELPHES

Domaine : Commerce, produits et marchés, vie des entreprises, gestion et environnement économique en France et en Europe. Base de données produites par le réseau de Chambres de Commerces et d'Industrie.

Données : issues de 900 revues économiques et professionnelles, et de journaux régionaux

### FRANCIS :

Domaine : Sciences humaines et sociales. Couverture mondiale

Données : issues de 1945 revues, dont près de 75% de l'Europe de l'Ouest.

Nous récapitulons les résultats de ces interrogations dans les tableaux suivants (<sup>4</sup>) :

- Tableau 1 : nombre d'occurrences de "competitive intelligence" pour les bases de données américaines les plus pertinentes.
- Tableau 2 : dates et sources des premières occurrences pour les termes "competitive intelligence", "business intelligence", "environmental scanning" dans les bases sélectionnées précédemment.
- Tableau 3 : même recherche pour les termes "intelligence économique", "veille technologique", "veille stratégique" dans la base DELPHES.
- Tableau 4 : date charnière à partir de laquelle l'emploi du terme augmente significativement pour les bases de la catégorie MULTIIND.
- Tableau 5 : occurrence des termes dans les bases en sciences sociales.

---

<sup>4</sup> Cette étude avait été présentée lors du séminaire CERSIDOC, ENSSIB VILLEURBANNE MAI 1996, intitulé "La veille en question".

### Tableau 1 : Nombre d'occurrences de "competitive intelligence"

Sélection des bases les plus pertinentes c'est à dire ayant le plus grand nombre de réponses à la requête "competitive intelligence"

ABI/INFORM	590
IAC PROMT	184
IAC TRADE &INDUSTRY DATABASE	585
TEXTLINE GLOBAL NEWS	130

Interrogation de la Catégorie MULTIIND (regroupement de bases concernant l'industrie) du serveur DIALOG. Date de l'interrogation : début 1996.

**Tableau 2 : Date et source des premières occurrences des termes : "competitive intelligence", "business intelligence", "environmental scanning"**

<b>ABI INFORM, PROMT, TEXTLINES, IAC, SOCIAL SCISEARCH, SOCIOLOGICAL ABSTRACTS, ECONOMIC LITTERATURE INDEX</b>		
<b>Terme</b>	<b>Première occurrence</b>	<b>Source et contenu de la première occurrence</b>
<b>Competitive intelligence (C.I)</b>	<b>1975 (2 références)</b>	<ul style="list-style-type: none"> <li>• <i>Public Relations Journal</i> "New standard for measuring public relations"</li> <li>• <i>Business Horizons</i> "Competitive business intelligence system"</li> </ul>
<b>Environmental scanning (E.S)</b>	<b>1967 13 références avant 1980</b>	Livre, New-York : Mc Millan "Scanning the business environment"
<b>Business Intelligence (B.I)</b>	<b>1971 (11 références avant 1980)</b>	<ul style="list-style-type: none"> <li>• <i>Industrial marketing management</i> "Intelligence : A technology for the 1980s ?" Une deuxième version de cet article en 1976 est donnée sous le nom : "Business Intelligence Systems : the need, development and integration"</li> <li>• <i>International Management</i> "Internal versus outside training"</li> </ul>

**Tableau 3 : Date et source des premières occurrences des termes "Intelligence économique", "Veille technologique", "Veille stratégique" dans la base française DELPHES**

<b>DELPHES (de l'origine à 1995)</b>			
<b>Terme</b>	<b>Nombre total d'occurrences</b>	<b>Date de la première occurrence</b>	<b>Source et contenu de la première occurrence</b>
<b>Intelligence économique</b>	<b>12</b>	<b>1992</b>	<ul style="list-style-type: none"> <li>• Le Moci (<i>Moniteur du commerce international</i>)</li> <li><b>"Etats-Unis : la surveillance offensive de la compétition"</b> <ul style="list-style-type: none"> <li>• Harbulot</li> </ul> </li> <li><b>"La machine de guerre économique"</b></li> </ul>
<b>Veille technologique</b>	<b>158</b>	<ul style="list-style-type: none"> <li>• <b>1982</b></li> <li>• <b>1983</b></li> </ul>	<ul style="list-style-type: none"> <li>• <i>Moniteur des travaux publics</i></li> <li><b>"L'observatoire technico-économique du bâtiment"</b></li> <li>• <i>L'expansion</i></li> <li><b>"Votre œil en Amérique"</b></li> </ul>
<b>Veille stratégique</b>		<b>1987</b>	<i>Le progrès technique</i> <b>"La veille de l'entreprise : un service qui n'a pas de prix"</b>

**Tableau 4 : Date charnière d'évolution des termes  
pour les bases issues de la catégorie MULTIIND**

<b>ABI INFORM, PROMT, TEXTLINES 1994-1995</b>	
<b><u>TERME</u></b>	<b><u>DATE CHARNIERE*</u></b>
<b>competitive intelligence</b>	<b>1987</b> 51 références avant 1987 785 de 87 à Mai 96
<b>Business intelligence</b>	<b>1990</b> 89 références avant 1990 1347 de 90 à Mai 96
<b>environmental scanning</b>	<b>augmentation régulière sans date charnière ; chute en 1996 (seulement 12 références en Mai 96)</b>

\* La date est la date à partir de laquelle l'utilisation du terme augmente très nettement sachant que le recensement a commencé en 1971.

**Commentaire des résultats :**

A partir de ces résultats, nous pouvons conclure par les trois points suivants :

1- Dès la fin des années 1960 se pose le problème de la surveillance de l'environnement de l'entreprise dans la littérature américaine. Ce n'est donc pas vraiment "*le contexte de mondialisation des échanges et d'émergence de la société de l'information*" [Livre vert, 1996], ni "*la fin de l'URSS et la modification profonde des enjeux stratégiques*" [Pichot-Duclos], "*la chute du mur de Berlin (...) la fin de l'affrontement bipolaire*" [Clerc, in Rapport mondial sur l'information, 1997] qui ont motivé l'apparition de cette question de la "veille" et de l'intelligence économique. La mondialisation est plus tardive, les bouleversements politiques liés à l'URSS également. En revanche, il est vrai que la "société de l'information" s'amorce.

2- On trouve quasiment simultanément, dès les années 1970, les termes "d'environnemental scanning" et de "competitive intelligence" (ou "business intelligence") dans les bases que nous avons consultées.

En revanche, en France le terme "d'intelligence économique" est plus tardif. Il apparaît seulement en 1992 dans DELPHES et dans FRANCIS :

- pour DELPHES il s'agit du Moci : "*Etats-unis : la surveillance offensive de la compétition : le point sur un nouveau concept américain...*" ;

- pour FRANCIS on cite l'ouvrage de Harbulot "*La machine de guerre économique : Etats-Unis, Japon, Europe*".

En France (cf. tableau 3), on a donc d'abord parlé de veille technologique, dix ans avant de parler d'intelligence économique, puis on a ensuite évoqué la notion de "veille stratégique".

**Au-delà des dates d'apparition de ces termes, ce qui ressort de la comparaison** entre les diverses bases de données (celles-ci pouvant être plus ou moins exhaustives dans leur sélection des références), **c'est que l'évolution du vocabulaire en France ne correspond pas à une évolution du vocabulaire américain** bien que le premier soit censé traduire le second. **Un choix a donc été opéré de la part des français : le thème de la veille/surveillance de l'environnement a été privilégié dans un premier temps à celui de l'intelligence économique alors que les deux concepts étaient "disponibles" en même temps (dès les années 1960)** <sup>5</sup>

3- Aux USA, les expressions "competitive intelligence" ou "business intelligence" sont nettement plus employées que celle "d'environnemental scanning" dans la littérature destinée à l'industrie ("business databases"). En France, nous avons vu le phénomène inverse : jusqu'à ces dernières années, on parlait de façon privilégiée de "veille".

Enfin, dans le vocabulaire des sciences sociales, c'est le terme "environmental scanning" qui reste apparemment le plus employé (Tableau 5).

---

<sup>5</sup> Les publications des années 1970 montrent que l'utilisation du terme est nouvelle mais courante.

**Tableau 5 : Nombre d'occurrences des termes "competitive intelligence",  
"business intelligence", et "environmental scanning"  
dans les bases de données en sciences sociales**

<b>SOCIAL SCI SEARCH, SOCIOLOGICAL ABSTRACTS, ECONOMIC LITERATURE INDEX (Interrogation en Juillet 1997)</b>	
<u><b>TERME</b></u>	<u><b>OCCURRENCES</b></u>
<b>competitive intelligence</b>	<b>41</b>
<b>Business intelligence</b>	<b>23</b>
<b>environmental scanning</b>	<b>79</b>

Que déduire de ces positionnements différents du vocabulaire alors même que l'un des termes se veut la traduction de l'autre (veille / environmental scanning ; intelligence économique / competitive intelligence) ? Correspondent-ils à une différence dans la façon d'aborder le problème ?

Examinons tout d'abord la présentation du problème aux Etats-Unis dans les années 1970.

#### I-1-2 APPARITION AUX USA DU THEME DE LA SURVEILLANCE DE L'ENVIRONNEMENT

Les premiers documents sur "environmental scanning system"/système de surveillance de l'environnement

L'un des pionniers en la matière est très probablement Aguilar qui, dans le cadre d'une thèse à la Graduate School of Business de l'université de Harvard, présente en 1966 une étude sur la façon dont le "top management" (les décideurs) acquiert de l'information pertinente sur les événements externes à l'entreprise afin de la guider dans ses orientations futures. Il appelle cette démarche "Scanning the business environment" ("scruter" l'environnement des affaires) : *"Scruter (Scanning) est l'activité d'acquérir de l'information. Ce livre traitera, non de toutes les activités de "surveillance" par les opérationnels, mais d'une partie déterminante de celles-ci- c'est-à-dire de la "surveillance" pour [obtenir] l'information sur les événements et les relations de l'entreprise avec son environnement, de la connaissance qui aiderait les*

*dirigeants dans leur tâche de prévision de l'action à mener dans le futur*"<sup>6</sup> [Aguilar 1967, p. 1. Notre traduction].

L'objectif d'Aguilar est alors de comprendre avec quel type d'information s'élabore la stratégie ; or celle-ci s'élabore au niveau des décideurs ("top management"). C'est donc à ce niveau de l'organisation que se situe le problème de la surveillance de l'environnement. Tel est d'ailleurs le raisonnement tenu par toutes les études en matière de stratégie. Pour Mintzberg il s'agit même d'une de ses prémisses : *"la responsabilité du processus doit rester entre les mains du PDG : cette personne est LE stratège"* [Mintzberg, 1994, p. 94].

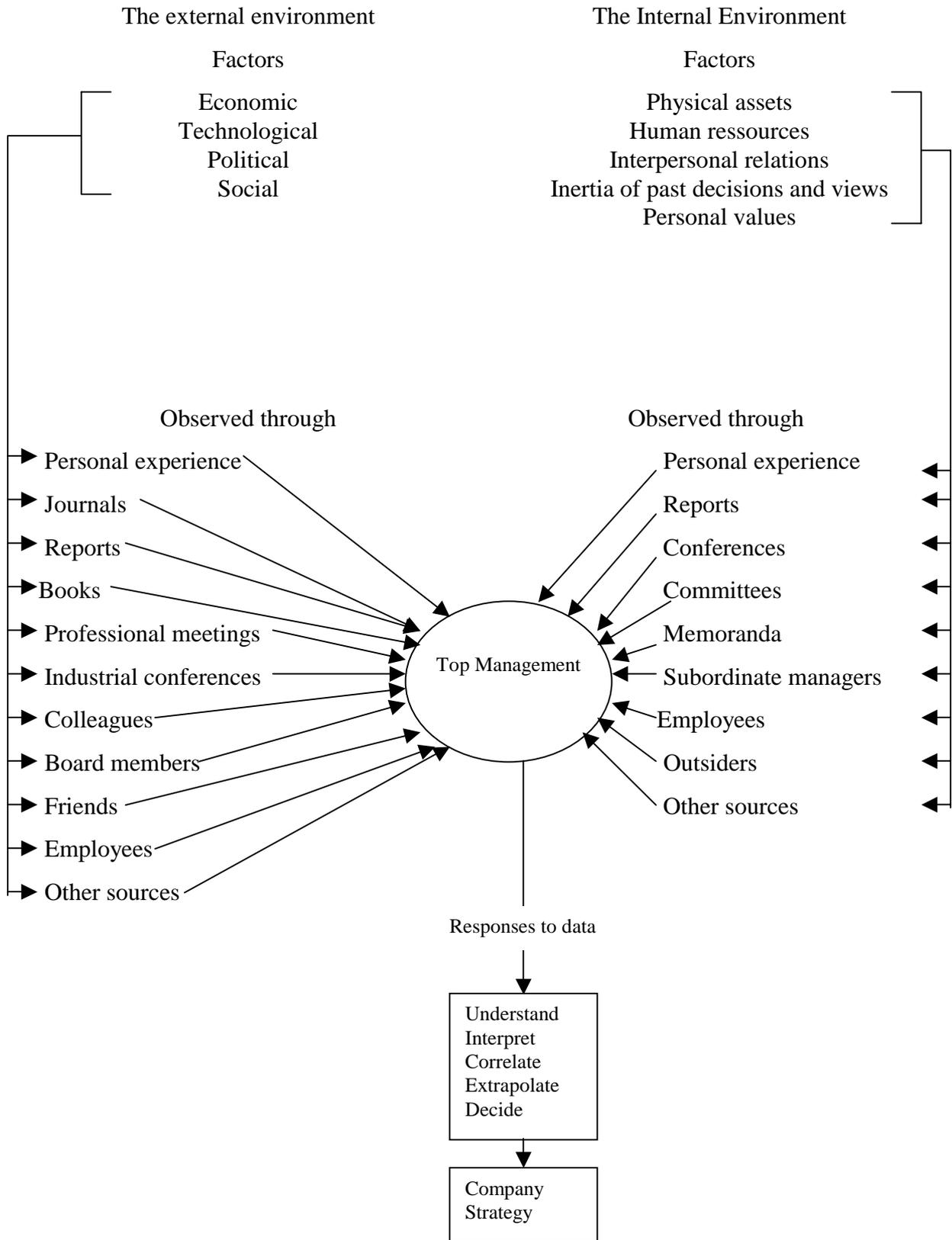
Ce n'est ni la prise de décision, ni l'analyse et l'interprétation de cette information, qui sont envisagées mais la manière dont l'information stratégique externe est utilisée.

Ce qu'Aguilar appelle "information externe" est celle qui relève de l'environnement extérieur à l'entreprise. Il s'agit des facteurs économiques, technologiques, politiques, sociaux par opposition à l'information provenant des ressources humaines, des relations interpersonnelles, de l'enseignement des décisions passées etc. Dans le schéma ci-dessous, "Some of the Many Sources for Strategic Information" (Quelques-unes des nombreuses sources pour l'information stratégique), l'auteur montre par quels "canaux" l'environnement externe et interne est observé par le "top management".

---

<sup>6</sup> *"Scanning is the activity of acquiring information. This book will be concerned, not with all scanning by executives, but with a critical part of the whole-that is, with scanning for information about events and relationships in a company's outside environment, the knowledge of which would assist top management in its task of charting the company's future course of action. The importance of such scanning derives from the importance of the decisions involved. Charting a company's future course of action i.e, formulating strategy and making long-range plans is unquestionably the chief executive's major responsibility and also his most difficult challenge"*

**Tableau 6 : Some of the Many Sources for Strategic Information**



D'après Aguilar [1967, p. 11].

Ce schéma, datant d'une trentaine d'années a été très largement répandu depuis les années 1980 pour justifier l'appel à la notion de veille. Un certain nombre d'éléments pointés dans le schéma (expérience personnelle, collègues, amis) font l'objet d'une distinction entre information formelle et informelle. Ce n'est pas cette distinction qui est utilisée mais celle existant entre environnement externe et environnement interne.

La détermination des domaines à surveiller, chez Aguilar, a été effectuée à partir de l'étude d'un échantillonnage d'entreprises de l'industrie chimique. Cette spécificité du terrain n'enlève rien, selon Aguilar, à une généralisation possible des résultats. Ces domaines sont récapitulés dans le tableau présenté "The Relationship of Industry Conditions to Areas of Scanning Attention", en Annexe I-B (p.62).

La tâche des "décideurs" est définie sous l'acronyme POSDCORD (Planning, Organizing, Staffing, Directing, Coordinating, Reporting, and Budgeting), ce qui est résumé dans l'expression "prise de décision". Par exemple, en ce qui concerne la planification de la diversification d'une ligne de produits, le management doit prendre des décisions sur des problèmes aussi complexes que les suivants :

- Quels objectifs doivent être recherchés pour de nouvelles entrées produit-marché ?
- Quelles diversifications alternatives paraissent prometteuses ?
- Comment ce choix doit-il être implémenté ?
- Quels contrôles doivent être effectués pour mesurer le progrès ?

Plus généralement, au-delà de cet ouvrage d'Aguilar, on peut dire que, dans la littérature américaine, la nécessité d'implanter un "strategic information scanning system" est lié aux problèmes de planification<sup>7</sup> et plus précisément, selon Aaker [1983], à ceux de la planification stratégique. Une menace ou une opportunité peut précipiter une décision stratégique ; et le calendrier de la planification risque de ne pas s'y adapter. D'où la nécessité de recourir à une analyse externe de l'environnement. Or ce type d'analyse, telle qu'elle est pratiquée dans le cadre de la planification, est jugé insatisfaisante car elle repose sur une collecte de l'information *ad hoc* et non systématisée.

Une quantité d'informations utiles sur l'environnement est perdue, dissipée ou inutilisée. A cela plusieurs raisons : la surveillance de l'environnement n'est pas dirigée, l'effort "de

---

<sup>7</sup> Nous expliquerons en I-2 (pp.34 et s.) à quels problèmes fait allusion ce discours et nous évoquerons brièvement la spécificité de la planification stratégique.

surveillance" (*"the scanning effort"*) n'est pas réparti correctement sur les participants, enfin cette surveillance ne repose pas sur des moyens de stockage, de recherche et de diffusion de l'information. Pour remédier à cette perte d'information il faut développer un système d'information stratégique (*"Strategic Information System"* ou SIS) en respectant les six étapes suivantes :

- spécification des besoins. Ces besoins concernent les concurrents, le marché (évolution des segments) et l'environnement. Ce dernier terme comprend tout ce qui est relatif aux technologies émergentes, à la réglementation gouvernementale, au développement économique national et régional ou aux tendances culturelles.
- spécification des sources : publications commerciales, clients, fournisseurs etc.
- identification des participants. Les participants du SIS sont ceux qui sont impliqués directement dans le processus de planification et ceux qui sont exposés à des sources d'information utiles (commerciaux, ingénieurs assistant à des congrès techniques etc.)
- assignation des tâches de surveillance ("scanning") à chacun.
- stockage et traitement : c'est le rôle de systèmes documentaires. Chacun doit savoir où envoyer l'information et la retrouver.
- diffusion

Tout ce travail doit immédiatement précéder la démarche de planification stratégique. C'est, en effet, le moment où le besoin d'information est le plus important.

Cette idée correspond à ce que nous appelons la "veille" en France. Toutefois il faut noter l'absence de référence à un quelconque modèle d'analyse stratégique (comme en France celui de Porter) pour organiser la collecte de l'information. L'accent est mis seulement sur le potentiel d'informations stratégiques contenu dans la firme et la nécessité d'organiser un système pour en éviter la perte.

### I-1-3 APPARITION DU THEME INTELLIGENCE ECONOMIQUE

Les premiers documents sur "competitive/business intelligence" :

On voit apparaître dans la littérature américaine des années 1970 la notion "*d'intelligence system*" appliquée au management. Ainsi parle-t-on de "*Competitive Business Intelligence Systems (CBIS)*" ou seulement de "*Business Intelligence Systems (BIS)*" ou encore de "*Scientific and Technical Information (STI) System*" appliqué au management dont l'intelligence économique ("*competitive intelligence*") en est l'application majeure <sup>8</sup>.

#### **I-1-3.1 Information scientifique et technique et intelligence économique**

L'américain Wolek [1977] précise que *l'une des plus importantes utilisations des services d'information réside dans les tâches de management telles que l'intelligence économique, l'évaluation de programme et la planification et qu'il s'agit là d'un changement considérable par rapport au passé où les systèmes de STI étaient avant tout destinés aux chercheurs et ingénieurs*. Wolek insiste sur cette idée, ce qui nous semble intéressant dans la mesure où il établit une filiation entre information scientifique et technique (plus ancienne), et information pour le management. Il souligne ainsi une évolution du rôle de l'information dans l'entreprise : l'intelligence économique en est la manifestation. Il rappelle toutefois le rôle prépondérant de l'information scientifique et technique dans la "Recherche et Développement" en signalant que, selon la plupart des études, les innovations viennent de l'extérieur de la firme qui les exploitent. Les travaux d'Olsen [1979], de Miller [1971], de Carter et Williams [1957] établissent à leur tour une corrélation entre la capacité des firmes à l'innovation et à acquérir et diffuser de l'information. Néanmoins les scientifiques et les ingénieurs ont tendance à agir en fonction de ce qu'ils savent déjà : "Des études récentes sur l'utilisation du STI dans l'industrie indique que les scientifiques et les ingénieurs agissent dans le cadre des ressources d'information déjà détenues plutôt que de réfléchir à celles qu'ils pourraient obtenir"<sup>9</sup> [Wolek, 1977, p. 40. Notre traduction].

---

<sup>8</sup> "*This is an important change from the past when the STI system's primary users were bench scientists and engineers. Some of the most important uses of information services are in managerial tasks such as competitive intelligence, program evaluation and planning.*"

<sup>9</sup> "*Recent studies of the use of STI in industry indicate that scientists and engineers operate in the framework of the information resources that they already have rather than thinking about what they might have.*"

L'utilisation limitée des possibilités des systèmes de STI par les chercheurs et ingénieurs, en même temps que la nécessité de détecter et d'évaluer les innovations, oblige le management à se préoccuper de STI. Mais c'est surtout la maintenance de produit ("*product maintenance*"), et le contrôle de nouveaux produits (diriger et gérer les impacts des produits) qui obligent les managers à réfléchir aux ressources d'information en fonction de leur stratégie. L'utilisation des bases de données dans ce contexte est donc considérée comme essentielle par rapport à sa destination première (qui serait la valorisation de STI par, et pour la Recherche et Développement).

### **I-1-3.2 Intelligence économique et management stratégique**

Mais l'idée d'intelligence économique va au-delà de celle d'un détournement de l'information scientifique et technique vers le management. Elle reprend à son compte les tâches classiquement dévolues au système de STI à savoir la collecte, le traitement et la diffusion de l'information et elles les subordonne aux objectifs du management stratégique. Objectifs dont nous avons dit qu'ils étaient sous-tendus par le thème de la surveillance de l'environnement. **La particularité de ce type d'analyse (celle du management stratégique telle qu'elle inspire l'intelligence économique) est qu'elle a pour base le rôle des forces compétitives dans le succès de la firme.** Comme le notent Gibbons et Prescott : "L'ascendant de la competitive intelligence, à la fois dans sa formulation stratégique et dans sa mise en place, peut-être attribuée en premier lieu à l'attention croissante des directions aux rôles joués par les critères de concurrence pour déterminer le succès de l'entreprise"<sup>10</sup> [1996, p. 163. (Notre traduction)].

Nous reviendrons sur ce point ci-après mais il convient de noter dès à présent que l'on retrouve dès 1975 un discours très actuel sur les objectifs d'un Competitive Business Intelligence System (CBIS ou BIS, selon les auteurs). Il faut mettre en place un système d'information sur l'environnement concurrentiel afin d'assurer en temps voulu une information valable sur ses concurrents principaux, déterminer la manière dont les concurrents peuvent affecter les intérêts courants de l'organisation, s'informer continuellement sur le système technologique, politique, économique, légal, social affectant la position concurrentielle. On

parlerait aujourd'hui, en France, de "veille globale". Trois catégories d'informations doivent être rassemblées : celles concernant le marketing, celles relatives à la production et au produit, celles touchant au domaine organisationnel et financier<sup>11</sup>. Les prémisses d'un CBIS sont, selon Cleland et King [1975], les suivantes :

- 1/ Un CBIS est essentiel pour réussir dans la compétition sur un marché particulier.
- 2/ Les marchés aujourd'hui changent de façon si considérable que les moyens informels pour maintenir la surveillance de la compétition sont inadéquats.
- 3/ Un "système total" est requis pour effectuer la fonction "intelligence". "Système total" signifie d'une part la centralisation de la fonction, d'autre part le recours à l'automatisation de tâches que nous qualifierions de "documentaires" (recherche d'information, diffusion de l'information)
- 4/ un BIS doit être hautement personnalisé. On trouve notamment ici exprimée l'idée d'une diffusion automatisée de l'information selon les profils d'utilisateurs, idée qui continue à motiver de nombreux travaux aujourd'hui en sciences de l'information.
- 5/ un tel système doit être orienté vers l'action. Il ne doit pas simplement produire des rapports de données agrégées mais indiquer l'action à privilégier.
- 6/ La "Business Intelligence" doit être obtenue à partir d'une variété de sources.
- 7/ un CBIS doit inclure des moyens de sécurité et de "contre-intelligence" sans recourir à des procédés illégaux.
- 8/ Un tel système n'a pas besoin de reposer sur des techniques illégales de collectes de données. On retrouve ici le thème, encore présent dans la littérature d'aujourd'hui, "d'intelligence systems" alors qu'on s'attache à montrer que celui d'espionnage est finalement contre-productif.

En conclusion, le développement de stratégies organisationnelles repose sur l'analyse des forces concurrentielles, des faiblesses et des actions probables. L'efficacité de l'analyse de la concurrence dépend plus de la logique et de la rigueur avec lesquelles les ressources d'information de l'entreprise sont organisées que de moyens ou de "subterfuges non éthiques" (notre traduction). Un CBIS doit aider à répondre à la question : quelle est la stratégie des concurrents et comment menace-t-elle notre position actuelle ou désirée sur le marché ?

---

<sup>10</sup> *"The ascendancy of competitive intelligence, in both strategy formulation and implementation, can be attributed primary to an increasing managerial awareness of the role that competitive forces play in determining firm success."*

<sup>11</sup> Voir annexe I-A

### I-1-3.3 Intelligence économique et organisation

- L'information pour la décision :

Jusque là on pouvait encore assimiler l'intelligence économique à ce que les Français ont appelé "veille" et considérer comme équivalents les termes "environmental scanning system" et "competitive intelligence systems". Les deux concepts sont en effet liés par ce domaine nouvellement ouvert à la réflexion scientifique dans les années 1960 aux Etats-Unis qu'est la stratégie des entreprises. Mais la notion "d'intelligence", au sens anglo-saxon du terme, implique une conception spécifique de l'information : un "intelligence system" n'est pas seulement un système d'information destiné à surveiller l'environnement. C'est un système qui cherche à informer les décideurs et implique, de ce fait, une conception de l'organisation. C'est ce thème de l'information au service de la décision qui justifie l'appel à la notion d'intelligence.

Le choix du terme "d'intelligence" se justifie, selon Pearce [1976]<sup>12</sup>, pour des raisons à la fois positives et négatives :

- Il s'agit de le distinguer à la fois des "data" (considérés comme des composants de l'information) et de "l'information" qui représente les données (data) organisées, traitées et structurées. La documentation (dispositif pour répondre au besoin d'information d'un utilisateur) se situe au niveau de l'information.
- "L'intelligence" se particularise en ce qu'elle implique le processus de décision lui-même, processus par lequel l'information est intégrée en termes de système à une organisation spécifique et des managers spécifiques"<sup>13</sup> [Pearce, 1976, p. 118. Notre traduction]

Une certaine confusion entoure le terme "d'information" : il renvoie soit à l'informatique ("*computer working*") soit à la documentation (Science de l'Information). En revanche, le terme "d'intelligence" a l'avantage de rappeler une analogie (même si elle est à utiliser prudemment) avec les activités traditionnelles de diplomatie et de bataille. En tant que tel il pourra désigner l'information telle qu'elle alimente le processus de décision, c'est-à-dire telle qu'elle détermine l'action. Fuld [1995] revient sur cette distinction entre "data", "information" et "intelligence" en donnant les exemples repris dans le tableau suivant :

---

<sup>12</sup> Il est intéressant de noter que cet article de Pearce [1976] a été publié dans sa première version en 1971 dans la revue "Industrial Marketing Management" sous le nom "Intelligence : A technology for the 1980s?"

<sup>13</sup> "*Intelligence moves us into the decision-process itself, whereby information is integrated in system terms applicable to specific organisations and specific managers.*"

**Tableau 7 : Distinction entre les termes "data", "information", "intelligence"**

DEFINITION	EXEMPLE
<p><i>DATA</i>. Bits dispersés et parcelles de connaissances;  <i>"Scattered bits and pieces of knowledge"</i></p>	<p>1990 : "Le rapport de Dun &amp; Bradstreet nous a révélé que l'usine de notre concurrent avait 100 employés".                      1993 : " L'un de nos commerciaux est passé dans l'entreprise de notre concurrent et a repéré seulement 30 voitures sur le parking."</p>
<p><i>INFORMATION</i>. Mise en commun de ces bits de connaissances.  <i>"A pooling of these bits of knowledge"</i></p>	<p>"Si on se base sur le rapport de Dun &amp; Bradstreet et sur les rapports de ventes, on constate que le concurrent a perdu des affaires."</p>
<p><i>ANALYSIS</i>. Information distillée.  <i>"Distilled information"</i></p>	<p>"Après avoir rassemblé plus d'information opérationnelle, et après avoir conduit une analyse comparative des pertes et des profits, il apparaît que le concurrent est devenu très performant. Il est au-delà des standards de l'industrie, et est devenu le meilleur".</p>
<p><i>INTELLIGENCE</i>. Ce qui permettra de prendre une décision.</p>	<p>" Le concurrent apparaît comme un bon candidat au rachat. Sa structure allégée et économe s'ajusterait parfaitement à nos opérations courantes".</p>

D'après Fuld [1995, p. 24]

- Le terme "intelligence" implique également l'idée de **système** :

*"Le terme d'intelligence indique que sommes attentifs à ce qui à trait au système : au feed-back, au contrôle, à la pertinence des actions, à l'action à mettre en place immédiatement, au management de l'unité, à la stratégie et à la politique."*<sup>14</sup> [Pearce, 1976, p. 119. Notre traduction].

Or cette notion de système, qui n'est pas ici appelée "système d'information", car cela serait redondant avec la notion d'intelligence, renvoie à l'idée d'une organisation des flux d'information. Il s'agit alors de gérer les mises à jour, la diffusion sélective et l'intégration aux circuits de décision, d'où les notions de feed-back, de pertinence, de débouché sur l'action etc. Depuis la fin des années 1950, une pluralité de techniques, de professions, d'études voient le

<sup>14</sup> "The term "Intelligence" signals that we are concerned with system matters-feed-back, control, relevant thing to do, immediate action, the operation of the unit managed, and strategy and policy."

jour afin d'assister la décision en fournissant et en traitant l'information, depuis l'information de fonctionnement (comptabilité analytique par exemple) jusqu'aux études stratégiques. Le développement des techniques macro-économiques de planification et de prévision (comptabilité nationale, planification indicative), des techniques quantitatives de prévision et d'études de marché en témoignent. Il est alors nécessaire pour le management de les filtrer et de les "router" vers les circuits décisionnels : *"On a atteint un seuil de taille et de complexité tel que le besoin se fait sentir d'avoir un filtre entre l'information, les procédés de manipulation et la décision qui aiderait le management dans ses choix, son évaluation et l'appréciation continue de la façon dont l'entreprise se transforme dans un environnement changeant."* [Pearce, 1976].

C'est l'organisation des flux d'information vers les niveaux décisionnels supérieurs ("top management") et leur convergence vers l'objectif de la prise de décision stratégique qui impliquent l'idée de système. Un *"intelligence service"* est donc aussi un *"intelligence system"*, ce que montre l'expérience militaire. L'armée en est un exemple : une stratégie, un plan d'action et une hiérarchie déterminent la circulation de l'information. La stratégie, par définition, est militaire : de stratos (armée) et agein (agir). Elle se définit comme : *"l'art de faire évoluer une armée sur un théâtre d'opérations jusqu'au moment où elle entre en contact avec l'ennemi."* [Le Petit Robert]. Elle s'oppose à la tactique qui représente l'ensemble des moyens mis en œuvre pour réaliser la stratégie.

Pourtant tout semble opposer le monde militaire et le monde économique. Comme le note le Général Fievet : *"Le monde militaire est caractérisé par son homogénéité et sa hiérarchie. Très organisé, structuré, discipliné, chaque échelon reçoit de l'échelon supérieur une mission qu'il ne discute pas. Ces personnels partagent un idéal commun et sont formés selon des normes et des procédés identiques. Le monde économique est particulièrement diversifié. Chaque entité économique réagit avec sa culture et ses habitudes propres. Chaque responsable d'entreprise se fixe lui-même sa propre mission. Le moteur du monde économique est constitué essentiellement par le profit."* [1992]

Alors que l'idée même d'intelligence économique s'inspire du modèle militaire autoritaire et hiérarchique (comme paradigme d'une circulation de l'information pensée pour alimenter l'action, et l'action stratégique en particulier) ; celui-ci est fustigé au nom de l'idée selon laquelle l'intelligence économique réclame un système en réseau. Dans ce dernier cas, le

processus de décision doit alors être décentralisé et coordonné. Selon Pearce [1976], traitant de l'organisation d'un "Intelligence Service" (dans le cadre d'un BIS) : *"Nous ne sommes pas intéressés ici par l'organisation, dans le sens d'une relation formelle entre personnes, rôles et fonctions, mais par des réseaux autonomes s'ajustant et s'adaptant, et pour lesquels le feedback est essentiel. La coopération est nécessaire pour atteindre les objectifs explicites ou tacites. De tels réseaux ont de nombreux centres de décision, d'importance et d'autorité variées, et engagent des processus qui, d'une certaine manière perçoivent, catégorisent, conservent, fournissent l'information, permettent d'arriver à prendre une décision, communiquent et réalisent l'action"*<sup>15</sup>.

Cette idée de réseau correspond au projet d'**Internet** qui voit le jour à cette période (en même temps que l'intelligence économique), paradoxalement (si l'on en juge l'opposition entre monde militaire et monde économique prônée par certains) en réponse à une commande militaire. La commande militaire consistait en un réseau non centralisé de telle sorte que si une partie du réseau était inopérationnelle, le réseau entier pouvait subsister.

Les mêmes préconisations sont encore avancées aujourd'hui en matière d'intelligence économique. Ainsi Gibbons et Prescott [1996] préconisent-ils des "processus parallèles d'intelligence économique" ("*Parallel competitive intelligence processes in organisations*"), l'intelligence économique ne devant pas relever d'une fonction unique mais d'un réseau de "nœuds" au sein de l'organisation. L'intelligence économique a donc participé à la promotion de la notion de réseau, paradigme de la société post-industrielle, selon la terminologie de Bell [1973], et dont la spécificité est d'être une "société de l'information". Ce principe général selon lequel une organisation basée sur l'information bouleverserait l'organisation industrielle traditionnelle est déjà résumée par Drucker en 1956 : *"Faire de l'autorité et de la responsabilité les principes d'organisation peut très bien se révéler la mauvaise façon de procéder lorsqu'il s'agit de personnes ayant un niveau d'étude élevé. Nous devons apprendre à fonctionner non pas dans un système fondé sur l'autorité et la responsabilité, c'est à dire un*

---

<sup>15</sup> "We are not concerned here with "organisation" meaning the formal relationships between people, roles and functions but with autonomous, self-directing, adjusting and adapting networks to which feed-back is critical. Co-operation is necessary to achieve stated or tacit goals. Such networks have many decision-making points of varying responsibility and authority and involve processes that in some sense perceive, categorise, store, retrieve, come to a decision, communicate and execute."

*ystème de commande, mais dans un système d'information et de décision...un système de jugement (discernement), de connaissances et d'attentes.*"<sup>16</sup> [Drucker, cité par Pearce, 1976. Notre traduction].

Un jugement concernant le changement des processus de contrôle nécessaires à la société de l'information est donc prôné au niveau de la gestion de la firme. Après le contrôle par la planification stratégique, c'est le contrôle par le système d'information qui est envisagé.

Selon le rapport Scott Morton [1995, p. 14-15]<sup>17</sup>, les technologies de l'information affectent le travail selon trois dimensions :

- le travail de production : robotique, contrôle de processus, capteurs.
- le travail de coordination : disparition de la distance dès lors que l'on s'intéresse aux flux d'information (le lieu de travail et les partenaires potentiels peuvent être réexaminés) ; annulation du temps (organisations situées dans des fuseaux horaires différents devant travailler ensemble, comme par exemple les systèmes de réservation des compagnies aériennes) ; mise à jour de la mémoire de l'organisation par une base de données commune ; travail de groupe.
- le travail de management : les tâches de direction (telles qu'elles sont liées à la perception de l'environnement externe : planification, système d'information sur la satisfaction des clients) et de contrôle (mesure des performances de l'organisation, interprétation de ces mesures par rapport au plan).

On peut dire que l'intelligence économique est le témoignage d'une prise de conscience, dès les années 1970, de cette modification du travail du management.

**Le terme "d'intelligence" implique donc à l'origine une conception de l'organisation,** c'est-à-dire de la façon dont les personnes, les matériels sont disposés, combinés, hiérarchisés et dont s'effectuent le partage des tâches, des attributions, et leur coordination. Il ne s'agit pas

---

<sup>16</sup> *"For the organisation of highly educated people authority and responsibility may well be the wrong principles of organisation... We will have to learn not a system of authority and responsibility- a system of command- but an information and decision system...one of judgement, knowledge and expectations."*

<sup>17</sup> Il s'agit d'un programme de recherche prospective initié en 1984 par la Sloan Management School sur l'entreprise des années 1990.

uniquement d'une théorie de la stratégie industrielle<sup>18</sup>. Organisation, information et stratégie seront donc les trois aspects constitutifs de l'intelligence économique.

Il est très frappant de constater que **dans ces premiers textes sur l'intelligence économique on ne voit pas encore apparaître la dimension politique du problème, alors que celle-ci est constitutive du concept d'intelligence économique en France**. C'est que cette dimension n'est pas liée à une réflexion sur la constitution d'un système national d'information comme c'est le cas en France. On reste au niveau de la firme ou de l'industrie dans les premiers textes américains comme d'ailleurs dans de nombreux textes sur ce sujet aujourd'hui, ce qui ne signifie pas, loin s'en faut qu'il n'y ait pas une politique en la matière et un rôle puissant de l'Etat fédéral. L'aboutissement de cette réflexion précoce sur le rôle de l'information dans la planification stratégique, alliée au développement des technologies de l'information, va plutôt conduire les américains à promouvoir une nouvelle guerre, la guerre économique, dont l'une des caractéristiques est d'être basée sur l'information.

Le travail des années 1965-1975 sur le rôle de l'information dans la gestion représente les prémisses d'un problème d'une envergure plus importante : le rôle appelé à jouer par les Etats-Unis dans la compétitivité mondiale grâce à l'information.

#### I-1-4 APPARITION DE CES NOTIONS EN FRANCE Premières références françaises

##### **I-1-4.1 La veille**

Dès le début des années 1980 apparaît dans la presse française le thème de la "veille". Dans un article consacré à l'expérience d'industriels français en Amérique, un journaliste de l'Expansion [Expansion 3/16 Juin 1983] explique que la modestie globale des performances françaises en Amérique du Nord "commence en amont dans un domaine pourtant essentiel aux techniques de pointe, la veille technologique et commerciale" **dont le Japon est l'exemple**. Dans ce domaine, il est indiqué que les français "*brillent par leur absence*" à quelques exceptions près. Parmi ces dernières, la filiale de la société Sofinnova à San-Francisco qui utilise la veille technologique comme un moyen pour identifier sa clientèle, l'Agence de l'informatique qui entretient un chargé de mission à Sunnyvale avec pour unique

---

<sup>18</sup> Nous reprenons ici une définition de l'organisation de Alberto et Combemale qui a le mérite de situer simplement le problème

fonction "la veille technologique". Au cœur de la Silicon Valley la tâche de ce chargé de mission consiste à dépouiller la presse spécialisée, les lettres d'information à diffusion restreinte, à suivre des colloques, des séminaires et des conférences, et surtout à entretenir un réseau de relations personnelles dans les laboratoires et les sociétés de la région, à l'université de Standford etc. Il travaille aussi bien sur des demandes ponctuelles que sur des enquêtes sectorielles, envoyant régulièrement des télex à Paris. Qu'une personne soit ainsi spécialisée pour accomplir de ces fonctions paraît très surprenant en 1983 : *"Jusqu'à une période toute récente, les Français ne croyaient pas à ce genre d'activité. Je passais pour un type bizarre qui avait trouvé la bonne planque en Californie"*.

La rentabilité de ce genre d'activité est mise en question, mais, en même temps, on remarque que les entreprises investissent des sommes importantes en Recherche et Développement sur des produits déjà dépassés par la concurrence.

Le bilan de la veille technologique et commerciale tel qu'il apparaît dans le discours de la presse fait apparaître l'hésitation du milieu industriel français face à une idée importée mais bien implantée en France au début des années 1980. Notons que les pouvoirs publics avaient déjà réagi à cette idée en assignant par exemple au Centre de Prospective et d'Evaluation des missions de veille. Celle ci était réalisée à partir de publications servant d'alerte. Une pluralité d'acteurs institutionnels ont ainsi été chargés de veille en France.

Les publications théoriques sur le sujet remontent au milieu des années 1980. Les travaux de Lesca constituent une reconnaissance de la notion de veille.

#### **I-1-4.2 L'intelligence économique**

Jusqu'au début des années 1990 on ne parle pas, en France, à notre connaissance, d'intelligence économique. Aux Etats-Unis les deux notions suivent parallèlement leur évolution. **L'exemple américain prévaut sur celui du Japon : l'intelligence économique est donc avant tout une notion américaine.**

En France, on la lie directement à la veille à laquelle elle vient se substituer. En passant de l'exemple japonais à celui des Etats-Unis on cherche à mettre l'accent sur *une "surveillance offensive de la compétition"* [MOCI du 12/12/92] ou encore à ce que Harbulot [1992] appelle *"la machine de guerre économique"* :

*"La fonction de Veille a été bien utile aux Français pour qu'ils s'intéressent à l'Intelligence. Mais l'Intelligence est offensive. C'est l'information évaluée, interprétée, utilisée par l'entreprise."*[Baumard, in Le Moci, op.cit]

Il s'agit alors de faire de l'information le nerf d'une nouvelle guerre : la guerre économique. La veille est jugée trop passive en ce sens qu'elle se limite à la collecte d'informations or "*La question n'est plus aujourd'hui de savoir combien d'ogives, quelle nouvelle arme ou quel secret scientifique et militaire il faut dérober, mais quels investissements réaliseront Hitachi, Siemens ou encore ITT dans les dix prochaines années, avec quel marketing, quelle innovation et quels partenaires.*"[*ibid*].

#### I-1-5 RESULTATS DE CETTE COMPARAISON DE L'USAGE DES TERMES EN FRANCE ET AUX USA :

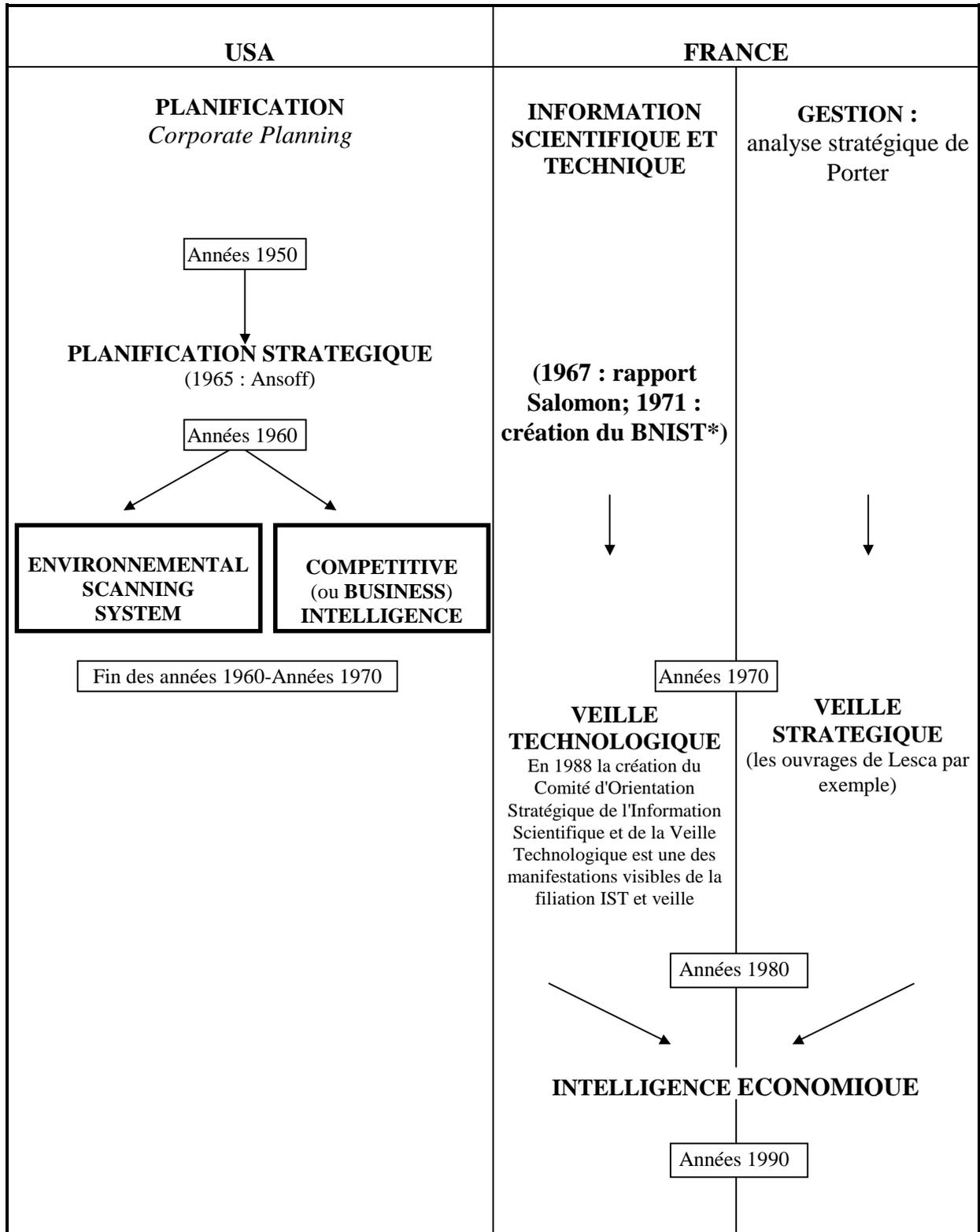
L'intelligence économique est un problème rémanent **depuis près de trente ans** aux Etats Unis. Les éléments caractéristiques de la notion "d'intelligence économique" en France (voir infra Introduction), à savoir le rôle de l'Etat, la coordination des acteurs pouvant jouer un rôle pour informer le milieu économique, les entreprises en particulier, n'est pas à l'origine de la notion d'intelligence économique telle que les américains, tenants des "*intelligence systems*", l'ont présentée dans les années 1970.

Nous avons repéré que, en France, le rôle de l'organisation dans l'information stratégique apparaît plus tardivement qu'aux Etats Unis. Dans un premier temps c'est la "veille" qui est privilégiée. Aujourd'hui chacun s'accorde à considérer la veille comme un processus passif par rapport à celui de l'intelligence économique.

Le modèle japonais, cité comme référence pour la veille, est absent de la littérature américaine. Pour les américains les questions de surveillance de l'environnement sont liées aux débats concernant la planification. Le modèle industriel japonais a été très étudié par les américains, en particulier en ce qui concerne l'industrie automobile. Mais le rôle de l'information dans le système japonais ne semble pas leur avoir servi d'exemple pour définir "l'environmental scanning system".

Nous résumons la comparaison entre l'apparition de ces notions en France et aux Etats-Unis dans le tableau page suivante.

**Tableau 8 : comparaison de la notion "d'intelligence économique"  
en France et aux USA**



\*Bureau National d'Information Scientifique et Technique

## **I-2 ELUCIDATION DES CONCEPTS**

### **I-2-1 SURVEILLER L'ENVIRONNEMENT : LES MODELES THEORIQUES UTILISES**

L'analyse de l'environnement en termes de menaces et opportunités est une des composantes de la gestion stratégique, elle-même issue de l'école de la planification. C'est à ces courants que se rattache la veille : elle se donne comme objectif d'alimenter la conception de la stratégie par une surveillance de l'environnement. En fait, l'idée de surveillance de l'environnement semble liée en France à un modèle stratégique bien spécifique, celui de l'américain Porter dont la spécificité est d'avoir précisément pensé la stratégie en termes de réaction de la firme aux forces concurrentielles qui constituent son environnement.

Aux Etats-Unis le courant de "*l'environmental scanning*" est rattaché à la planification stratégique dont Ansoff est le grand initiateur. Ansoff a voulu adapter la planification à un environnement désormais qualifié de turbulent et de changeant, considérant que des systèmes dits "quasi-analytiques" [Ansoff, 1965] pouvaient programmer le changement et l'intégrer dans la stratégie. Il considère comme possible la surveillance des signaux faibles et une gestion de la surprise. Le mouvement de "*l'environmental scanning*" va se greffer sur ces idées. Ce faisant, ils mettront en évidence le paradoxe consistant à vouloir prévoir le changement alors qu'un bon dispositif d'information suffirait à alimenter la décision stratégique.

C'est avec les **remises en cause des méthodes traditionnelles de la planification** que le rôle de l'information dans la gestion, en particulier le problème de l'information des dirigeants, va être évoqué. La planification qui était avant tout une méthode budgétaire dans les années 1950 est progressivement devenue une méthode de travail, dite "PPBS" ou "Planning-Programming-Budgeting-System" "*véritable obsession des entreprises américaines et du gouvernement américain*" selon Mintzberg [1994].

La planification traditionnelle consiste en une démarche en trois étapes : l'élaboration d'objectifs, la mise au point de programmes, l'établissement de budgets. Les performances sont analysées en termes d'écart par le contrôleur budgétaire. Ce dernier vérifie ainsi dans quelle mesure le plan a été réalisé. Dénoncée pour sa lourdeur bureaucratique (participant à ce que Crozier [1963] a appelé le "phénomène bureaucratique"), elle a également été critiquée pour ne pas intégrer la dimension environnementale. Elle privilégiait la formalisation d'un

plan écrit, la hiérarchisation des budgets et des objectifs sur l'adaptation. Par la remise en cause de ces méthodes traditionnelles, on a cherché à transformer la planification en gestion du risque et de l'incertitude.

En soulignant que la planification ne peut plus être conçue dans un environnement stable mais complexe (comprenant plus d'acteurs) et turbulent (changements plus fréquents et plus profonds), on a mis l'accent sur la "quasi analyticit " voire la "non analyticit " des processus de d cision qui mettent en  uvre les proc dures du plan. Les informations sont entach es d'incertitude, d'o  le caract re heuristique des d cisions. On a recours   de multiples crit res qui peuvent ne pas  tre quantitatifs, pour trier les informations, choisir et d cider.

Des m thodes d'analyse strat gique vont alors se d velopper, tout comme l'analyse des processus de d cision strat gique.

Le d veloppement de la planification strat gique par Ansoff au cours des ann es 1960-70 (notamment Ansoff [1965]), ou du courant d cisionnel initi  par le prix Nobel d' conomie Simon t moignent de l'amorce de cette remise en cause.

La planification n'est pas l'apanage des entreprises. Elle est aussi le fait de la politique et des Etats.

Il faut noter   quel point ce probl me de la planification a des implications tr s actuelles en France sur le plan politique en particulier. On peut constater qu'avec la cr ation, en 1992, de l'ADIT (Agence pour la Diffusion de l'Information Technologique), administration   caract re industriel et commercial charg e de la politique fran aise en mati re d'intelligence  conomique, le Centre de Prospective et d'Evaluation (CPE) a disparu. L'ADIT a pour vocation de se substituer au CPE. On constate  galement que le Commissariat G n ral du Plan "*est r duit au ch mage technique*" [Le Monde 4/10/1997]. Quant au Centre d'Etudes, de Prospective et d'Informations Internationale (CEPII), organisme rattach  au Commissariat g n ral du Plan, il semble  tre rel gu    une place tr s secondaire. Plusieurs interpr tations de cet  tat de fait sont possibles, notamment celle accusant la politique publique<sup>19</sup> de se contenter du court terme. Il n'est du reste pas anodin que ce mouvement de d fiance vis   vis des organismes de planification et de prospective soit contemporain d'un engouement pour l'intelligence  conomique et de la mise en place d'une politique publique dans le domaine.

C'est donc bien avant l'analyse strat gique de Porter, que le mouvement de "*l'environmental scanning*" s'est d velopp . Il doit  tre rattach  selon nous   ce que Mintzberg [1994] a appel 

la "*grandeur et la décadence de la planification stratégique*". Nous montrerons dans les lignes qui suivent comment s'organisent ces modèles.

### **I-2-1.1 Le management stratégique et le thème de l'analyse de l'environnement**

La conception de l'environnement dans la pensée stratégique peut varier, d'où cette rapide typologie que nous proposons afin de repérer les modèles dont est issue la veille. Le management stratégique envisage la stratégie industrielle selon trois approches :

- l'approche "*corporate et business strategy*" : il s'agit essentiellement d'une **analyse par l'activité**. L'environnement est ici réduit au jeu du marché et sa structure est censée imposer largement à l'entreprise les limites de sa stratégie. On trouve ici toutes les techniques de choix de portefeuille.

- l'approche environnementale : elle repose sur un avantage relatif par rapport à ses concurrents à l'intérieur d'un secteur. On peut parler d'**analyse par les forces concurrentielles**. Selon l'analyse des forces concurrentielles, l'entreprise peut agir sur les structures du marché et les conditions de la concurrence. L'approche par les forces concurrentielles complexifie la notion d'environnement en élargissant le champ concurrentiel pris en compte. La stratégie est étudiée au niveau des structures industrielles, c'est à dire de l'industrie dans laquelle la firme évolue ou les industries avec lesquelles elle est en concurrence.

- l'approche par les processus de décision ou par la façon dont se prennent les décisions stratégiques. L'approche relève ici d'une **analyse par l'organisation**.

#### L'Analyse par l'activité

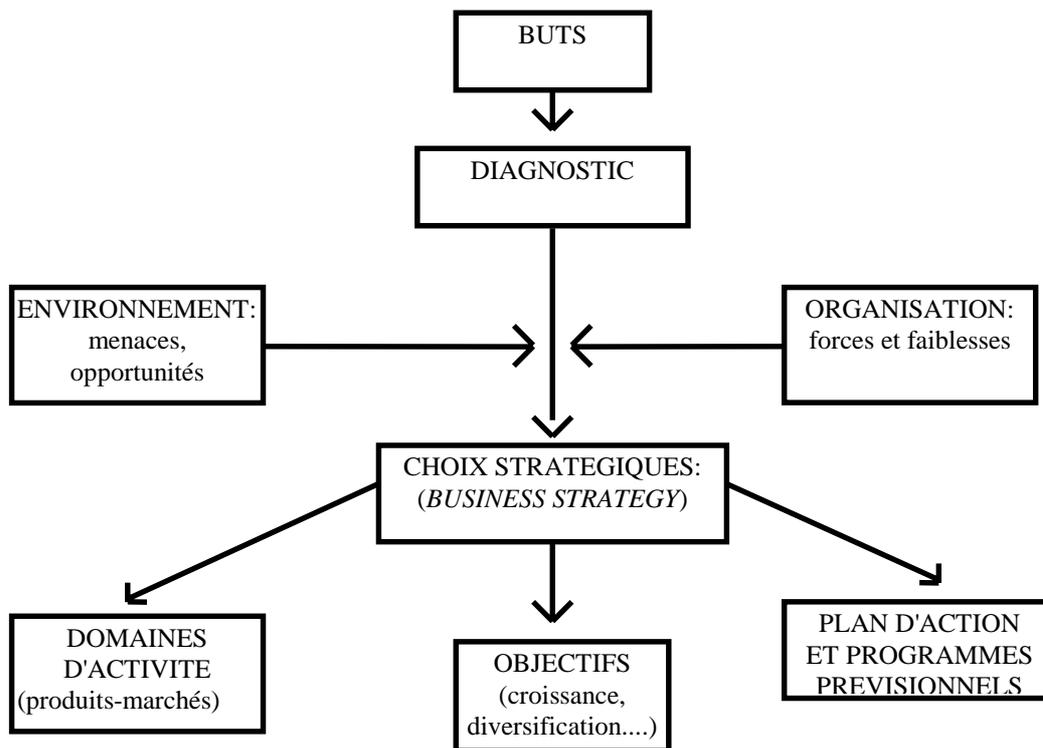
Elle relève de la "*business strategy*" dans la mesure où elle situe la stratégie au niveau de chaque division produit-marché, par opposition à la "*corporate strategy*" qui se situe au niveau de la politique générale de l'entreprise. L'objectif est d'articuler ces deux niveaux afin de maximiser la valeur des actions, et donc le profit, par la constitution d'un portefeuille. L'un des premiers modèles de planification stratégique s'appelait le modèle "SOWT" [Strength (force), Opportunities ("opportunités"), Weakness (faiblesse), Threat (menace)] car il mettait en

---

<sup>19</sup> comme celle des entreprises d'ailleurs : il n'y a qu'à constater le problème de l'investissement dans la recherche !

balance les forces et faiblesses de l'organisation et les menaces et opportunités détectables dans l'environnement selon le schéma général suivant :

**Tableau 9 : le modèle SOWT**



Dans ce modèle, l'environnement est vu comme une entité faite de menaces et d'opportunités que l'on détecte à partir de faits et d'observations quantifiées (bilans, parts de marchés,...) et auquel fait face l'organisation. L'environnement est assimilé au marché, lequel semble évoluer indépendamment de la stratégie des firmes, dirigé par la seule évolution des utilisateurs du produit.

La "*business strategy*" concerne surtout le marketing car elle touche essentiellement au développement des produits (biens et services) offerts par l'entreprise. L'activité s'exprime là en termes de "produits-marché" ou "technologies-produits-marchés".

- Les matrices d'analyse stratégique : la stratégie réduite aux techniques de choix de portefeuille

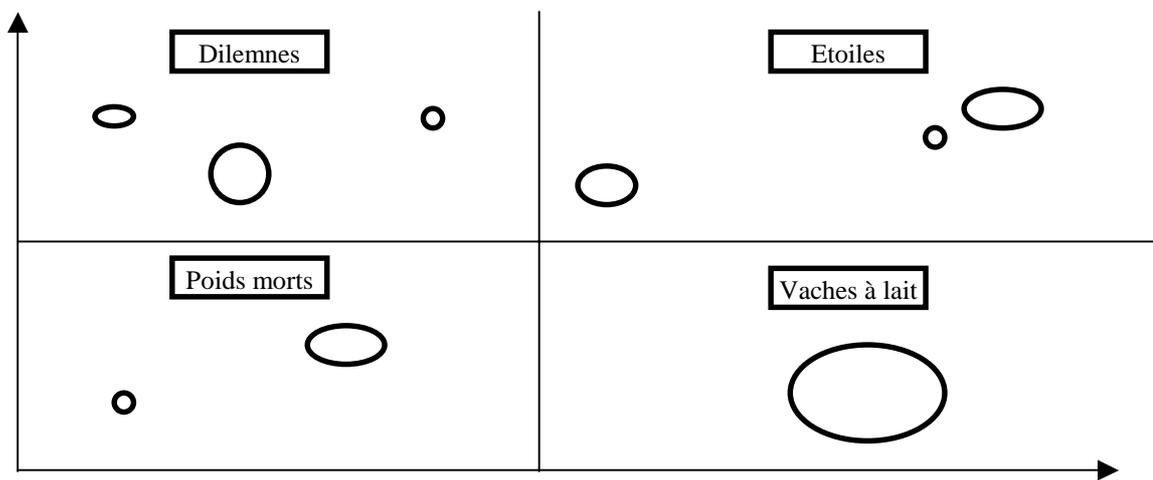
Dans ce domaine, se sont développées des matrices d'analyse stratégique émanant de cabinets de consultants comme le Boston Consulting Group (BCG), le cabinet MacKinsey, ou celui

d'Arthur D. Little (ADL), pour ne citer que les plus connues et les plus anciennes. Ce sont des matrices de ce type qui sont parfois utilisées comme méthode de mise en place de la veille (nous avons pu constater qu'elles s'enseignaient récemment sous le nom de "veille concurrentielle et commerciale" (cf. supra : *Introduction de la partie I*).

Dans sa première version, la méthode BCG analyse le portefeuille d'activités de l'entreprise selon deux critères : le taux de croissance du domaine d'activité et la part de marché relative (mesurée par le ratio : part de marché de l'entreprise / part de marché du concurrent principal) détenue par l'entreprise dans le domaine d'activité).

**Tableau 10 : matrice du BCG**

*Taux de croissance  
du domaine d'activité*



*Part de marché relative  
détenue par l'entreprise*

Les surfaces des cercles sont proportionnelles aux chiffres d'affaires réalisés par les différents domaines d'activité.

Ce modèle a été critiqué parce qu'il repose essentiellement sur la notion d'économie d'échelle<sup>20</sup>, et parce qu'il lie l'importance de la part de marché aux taux de rentabilité. Ce modèle se révèle être inadapté à l'époque de la recherche de séries courtes, de juste à temps et de flux continu. Une seconde version a été proposée. Cette seconde version croise deux critères :

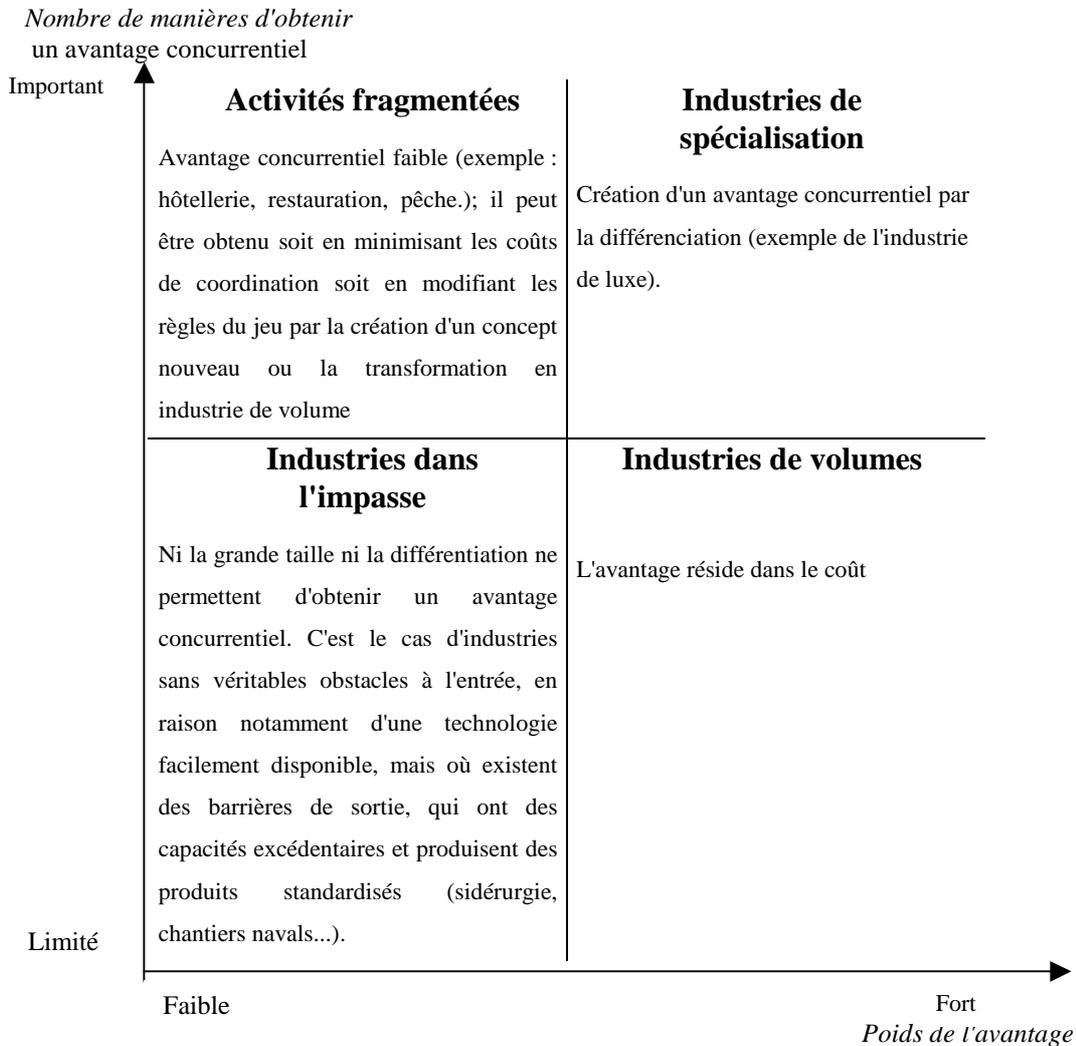
- le nombre de manières différentes pour obtenir un avantage concurrentiel (politique de gamme, coût, qualité, technologie, design, etc.). Ce nombre est faible, par exemple, dans les

<sup>20</sup> Cette notion exprime la baisse du coût unitaire d'un produit corrélativement à l'augmentation de sa production au sein de la firme.

industries produisant des biens de consommation standardisés (l'avantage repose essentiellement sur les coûts) et fort dans les services ou le commerce de détail.

- l'importance de l'avantage concurrentiel.

**Tableau 11 : Deuxième version de la matrice du BCG**



D'après Alberto et Combemale [1993]

La grille d'analyse MacKinsey met en relation la valeur de l'activité et la position concurrentielle, celle d'Arthur D.Little met en relation la phase de cycle de vie de l'activité et la position concurrentielle.

- La stratégie par les domaines d'activité ("strategic business units")

Cette approche intègre les ressources dans les choix stratégiques : le champ stratégique se définit en fonction des forces internes de l'entreprise et des opportunités externes en termes de produits-marchés. Toute acquisition, tout lancement de produit nouveau, ou encore toute diversification ne sont envisagés qu'en fonction des compétences distinctives spécifiques à l'entreprise. Ce sera par exemple, en Recherche-Développement, l'articulation de la "recherche-produit" aux études de marché et la maîtrise des tactiques de publicité et de promotion.

La notion de domaine d'activité repose sur le découpage des activités en segments c'est-à-dire en unités homogènes du point de vue de l'action stratégique à mener. Voici un exemple de ce type de découpage : *"Soit une entreprise fabriquant des aliments pour chiens et pour chats selon des technologies différentes (conserves, granulés, etc.) et pour des couches de clientèle différentes (donc des circuits de distributions différents : supermarchés, magasins spécialisés, etc.), avec des marques différentes, dans des pays différents. On voit qu'il s'agit de regrouper les activités en segments présentant fondamentalement les mêmes stratégies. Par exemple, on aura un premier segment stratégique correspondant à une stratégie de coûts bas, et de pénétration d'un marché de grande consommation, quel que soit le produit, mais avec la recherche de grandes séries (économies d'échelle), de gammes étroites (peu d'économie de variété), de circuits de distribution courts (centrales d'achat, grandes surfaces). Un autre segment comprendra les produits axés sur la différenciation (séries plus courtes, gammes larges, magasins spécialisés, etc.)."* [Exemple cité par Marchesnay, 1993]

Nous ne présenterons pas les différentes stratégies des domaines d'activité stratégique : nous souhaitons seulement ici distinguer les approches de la stratégie.

Cette attention aux compétences actuelles et potentielles de l'entreprise ("*value added skill*") est aujourd'hui centrale dans la gestion stratégique de l'information (voir I-3). La documentation tend vers une gestion des connaissances ("*knowledge management*") en tentant de fournir les outils nécessaires à la formalisation et la capitalisation des savoirs spécifiques d'une organisation.

## Analyse par les forces concurrentielles

Dans l'analyse des forces concurrentielles, on ne considère plus que les forces du marché déterminent le type de concurrence et que l'entreprise n'aurait qu'à s'y adapter. Au marché se substitue l'industrie (le secteur) au sein de laquelle l'entreprise peut modifier les conditions de son environnement.

La stratégie consiste, pour une firme placée dans un secteur, en des "actions offensives ou défensives qui visent à la mettre dans une situation tenable vis-à-vis des cinq forces concurrentielles" [Porter, 1982, p. 33]. Ces cinq forces sont présentées dans le tableau 12.

Quelle que soit la démarche choisie<sup>21</sup>, la firme disposera de trois stratégies fondamentales (ou génériques) : la domination au niveau des coûts (obtention de coûts faibles), la différenciation (créer quelque chose de ressenti comme unique par la clientèle), la concentration de l'activité (s'organiser autour d'une clientèle ciblée en étant plus efficace ou en consommant moins de ressources que les concurrents qui luttent dans un domaine plus large par exemple).

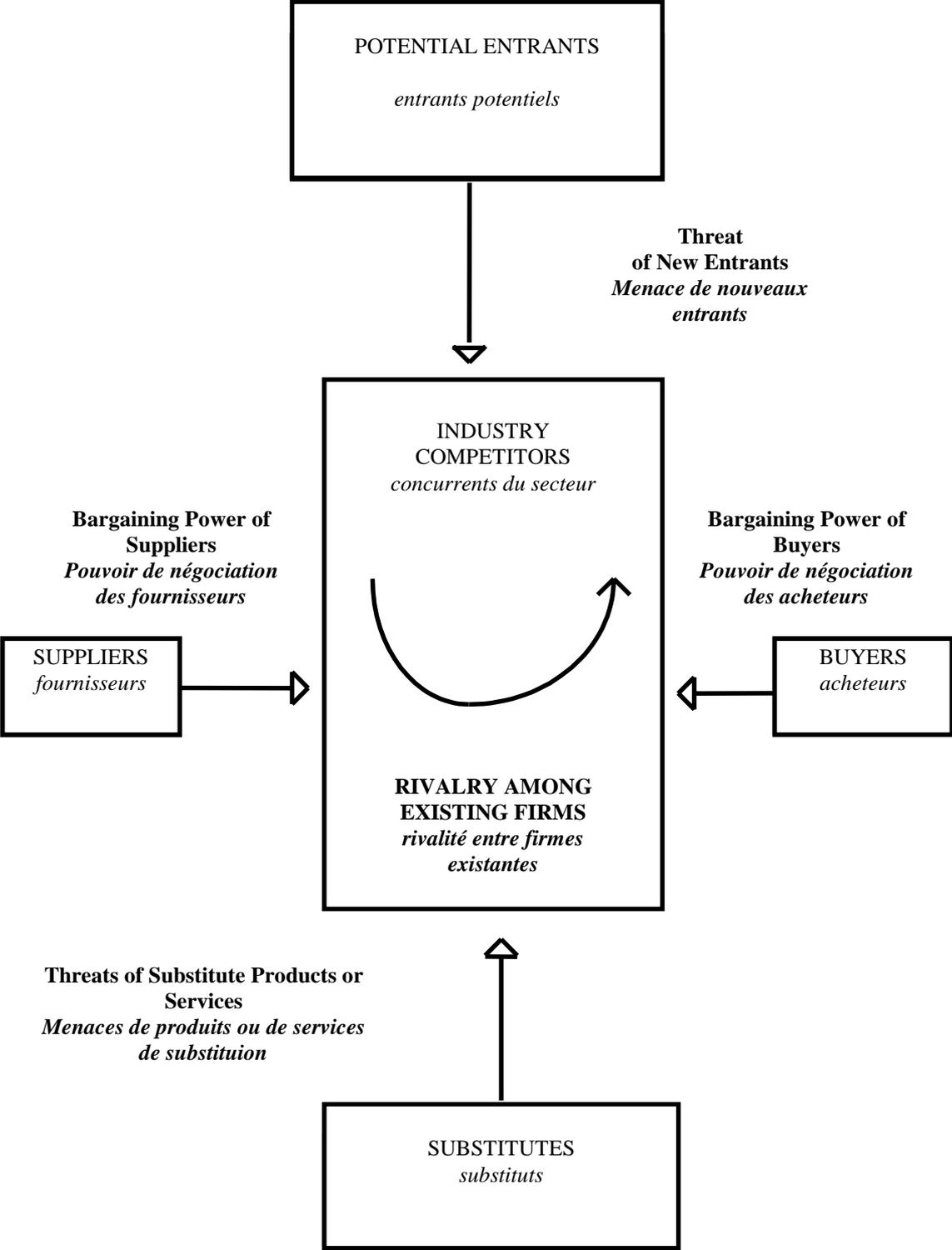
Grâce à ces stratégies fondamentales l'entreprise tente de subsister ou d'éliminer ses concurrents en élevant des barrières à la mobilité. L'industrie se modifie et se structure en différents groupes stratégiques qui se protègent les uns les autres par des barrières.

Ainsi l'on voit en quoi cette approche environnementale de la stratégie se distingue de l'approche précédente : une stratégie basée sur la dynamique concurrentielle se substitue à une stratégie focalisée sur les marchés finaux. Une certaine prise en compte du rôle de l'information en découle, comme nous le notons ci-après. C'est ce modèle de la stratégie qui a servi de référence à ceux qui ont introduit la notion de "veille" en France (cf. les deux schémas qui suivent).

---

<sup>21</sup> Assurer la meilleure défense possible contre le déplacement actuel des forces de la concurrence ou influencer sur l'équilibre des forces ou encore anticiper les modifications des facteurs qui régissent ces forces et réagir en exploitant cette évolution.

**Tableau 12 : les cinq forces qui déterminent la rentabilité de l'industrie**

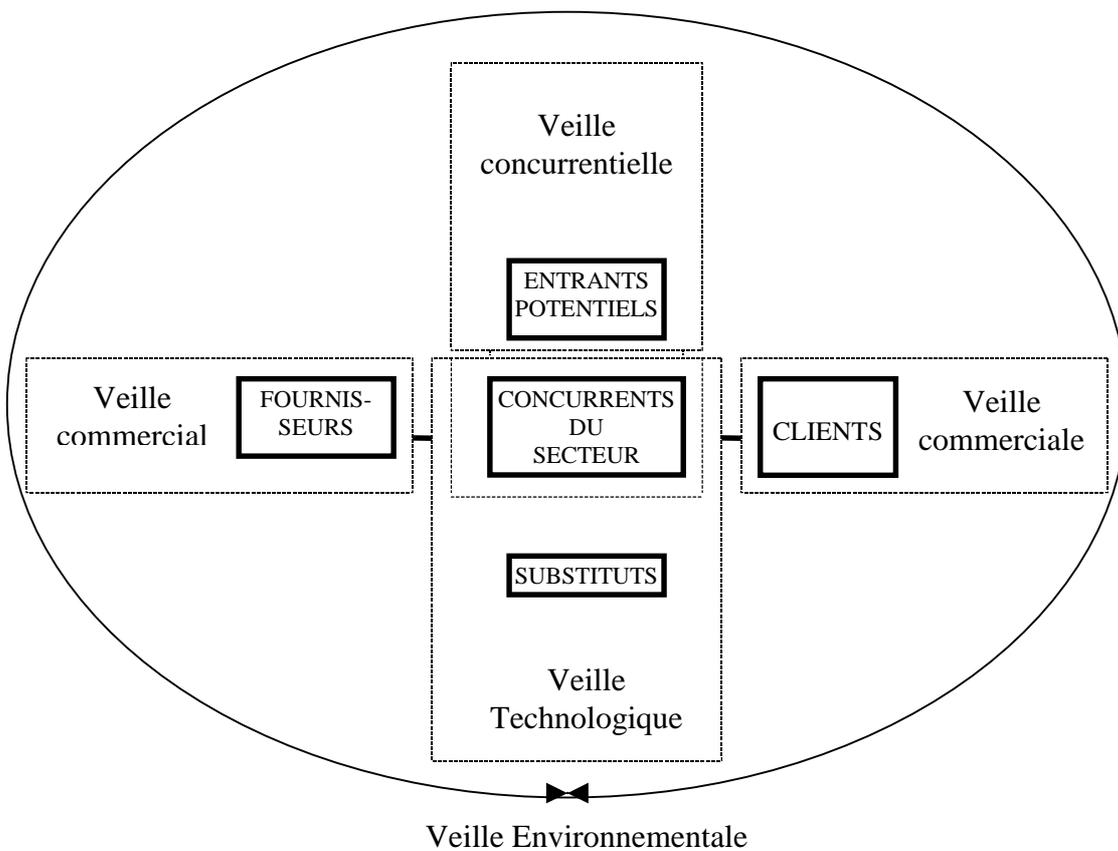


D'après Porter [1982, p. 5]  
*Les forces sont en caractère gras*

- Porter et la surveillance de l'environnement

Reprenant ce modèle de Porter, les français Martinet et Ribault [1989] construisent une théorie de la veille.

**Tableau 13 : Les cinq paramètres qui commandent la concurrence au sein d'un secteur et les quatre types de veille associés**

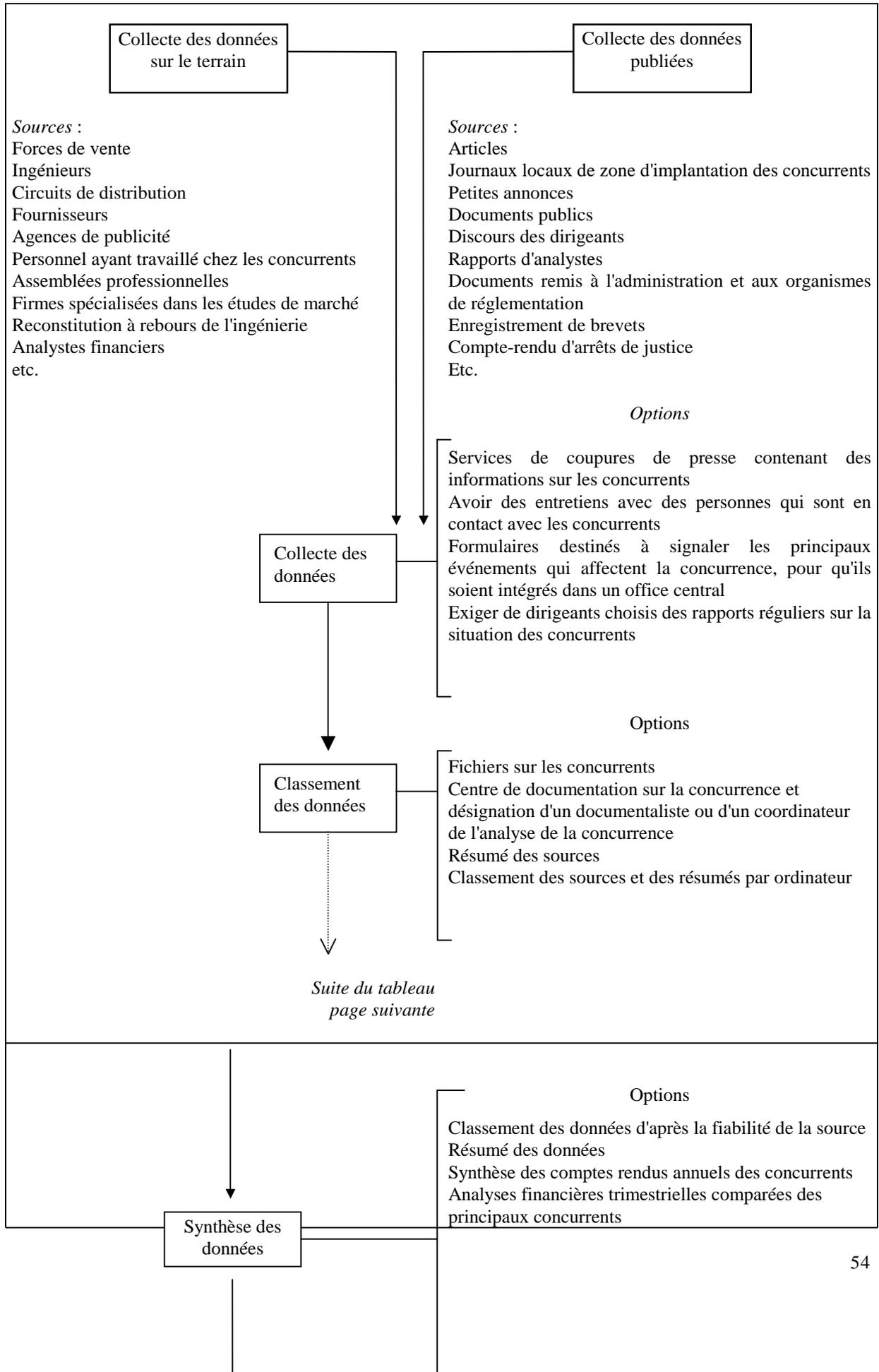


D'après Martinet et Ribault, 1989, p.24

Cette approche n'est pas précisément celle de Porter. Porter évoque, dans son ouvrage "Choix stratégiques et Concurrence" (traduction française de "*Competitive Strategy*"), la nécessité d'un système de *renseignement* sur la concurrence [1982, p.79], en particulier pour *prévoir* les conditions qui prévaudront plus tard dans le secteur. Porter insiste sur la nécessaire organisation d'un tel mécanisme : "*L'efficacité du processus exige pour être atteinte, un mécanisme organisé - une sorte de système de renseignement sur la concurrence.*" [ibid, p.79]

Porter représente sous la forme du diagramme suivant les fonctions d'un tel système.

**Tableau 14 : "Les fonctions d'un système de renseignement sur la concurrence"**



Analyse des gammes de produits relatives  
Evaluation des courbes de coût et des coûts relatifs des concurrents  
Evaluations financières en forme des concurrents dans le cadre de différents scénarios sur l'économie, les prix et les conditions de la concurrence

#### Options

Collecte régulière des coupures de presse à l'usage des dirigeants occupant des postes clés  
Lettre d'information ou rapports de situation réguliers du concurrent  
Rapports d'analyse en profondeur, sans cesse mis à jour, sur les concurrents  
Réunions, dans le cadre du processus de planification, spécialement consacrées à un exposé sur la concurrence

D'après Porter [1982, p.80-81]

Porter s'interroge également sur le type de données qu'il est nécessaire de recueillir pour analyser un secteur et la concurrence. Pour ce faire, il présente un ensemble de domaines pour lesquels il faut collecter des données brutes (ibid p.402 : "Type de données brutes pour une analyse sectorielle").

- Les limites du modèle de Porter

Porter a mis en lumière le rôle de l'environnement dans la stratégie industrielle. Il a été repris par le courant français de la veille comme modèle à la mise en place de celle-ci.

Toutefois, malgré le grand succès du modèle de Porter, celui-ci a été soumis à la critique<sup>22</sup>. Celles-ci n'apparaissent jamais dès lors qu'on l'utilise pour justifier la veille.

Parmi ces critiques, nous retiendrons en particulier les points suivants :

- Porter ne prend pas en compte les spécificités organisationnelles de chaque firme. Il s'intéresse à la stratégie au niveau des structures industrielles. Il lie donc la stratégie à son environnement, c'est-à-dire au contexte industriel dans lequel évolue la firme (que ce soit en synergie ou en concurrence). Par conséquent, l'analyse de la firme n'est pas orientée vers l'étude de la gestion et du développement des capacités propres, internes à la firme comme la technologie. La notion de capacité spécifique est introduite uniquement comme un moyen utilisé par la firme pour dissuader l'entrée des concurrents sur le marché. L'auteur n'explique ni sa formation ni sa construction.
- La technologie ne fait pas partie pour lui des capacités et ressources de l'entreprise ; elle constitue une activité de soutien aux activités principales appartenant à la chaîne de valeur.
- Porter ne prend en compte que le caractère de substituabilité d'une technologie (le traitement de texte remplace la machine à écrire) et jamais la complémentarité des technologies.
- La technologie est analysée seulement comme une arme stratégique, eu égard à la menace qu'elle fait peser sur les concurrents. Les technologies sont vues comme des phénomènes qui peuvent confisquer des effets d'expérience acquis par une industrie et donc anéantir la compétitivité. Les conditions d'acquisition et de développement ne sont pas prises en compte.

### Analyse par les processus de décision

Le problème du rôle de l'information dans la gestion est posé avec la planification stratégique, mais il est étudié en tant que tel par Simon. Ce prix Nobel d'économie, connu pour ses travaux novateurs en intelligence artificielle, a pour point de départ l'étude de l'automatisation des activités de gestion.

Nous prendrons ici comme travail de référence son ouvrage : *Le nouveau management : la décision par les ordinateurs* [1990]. Dans ce cadre (chapitre 4) il est, en effet, amené à définir le renseignement ("intelligence") par rapport à la stratégie, distinction qui éclaire, selon nous, la notion "d'intelligence économique".

---

<sup>22</sup> Voir par exemple la thèse de Hendricks [1995, pp. 31-54].

a) Simon étudie la prise de décision au sein du monde des affaires. Son projet est d'automatiser un plus grand nombre de décisions non programmées grâce à la programmation heuristique. Cela le conduit à l'analyse suivante des processus de prise de décision des dirigeants.

Les problèmes posés aux dirigeants à différents niveaux des organisations peuvent, selon Simon, être classés en fonction des critères suivants : sont-ils bien structurés, habituels et préétablis ? Il en est ainsi, par exemple, de l'acquisition des fournitures de bureau, ou la fixation des prix de produits standard. Relèvent-ils de décisions non programmées comme les décisions fondamentales (et définitives) de fabrication d'une nouvelle gamme de produits ou les stratégies de négociation destinées à préparer une nouvelle convention avec les travailleurs ? En général plus la position du dirigeant est importante au sein de l'organisation moins ses décisions sont "programmées". En moyenne, les décisions prises par le président et le vice-président sont en effet moins "programmées" que celles prises par le chef du département de la production ou par le directeur d'usine.

Simon distingue quatre phases du processus de prise de décision : **l'activité de renseignement, l'activité de conception, l'activité de sélection et celle de constatation.**

La première concerne **l'étude de l'environnement pour y rechercher des conditions qui appellent une décision. C'est le renseignement au sens militaire du terme.**

La seconde est constituée par l'invention, le développement et l'analyse des différents modes d'action possibles.

La troisième est relative au choix, parmi les actions possibles, d'un mode d'action particulier.

La constatation est une phase de bilan des décisions passées.

Pour illustrer ces phases Simon propose un exemple : celui de l'introduction des ordinateurs dans le processus de gestion et le développement progressif des systèmes informatiques, d'une part, dans le monde des affaires et, d'autre part, dans le secteur public.

b) A propos de l'échec des systèmes d'information de gestion, Simon oppose renseignement et planification stratégique.

L'essentiel des décisions pour lesquelles la haute direction joue un rôle capital nécessite des informations qui n'émanent pas de l'entreprise elle-même. Il semble que l'on assisterait à un mouvement inverse aujourd'hui : *"Les hauts dirigeants se trouvent au point de jonction de l'entreprise avec son environnement de clients, de concurrents, de sources de financement et de systèmes politiques et sociaux. Leur principal objectif est de découvrir comment l'entreprise peut vivre efficacement et de façon profitable dans cet environnement extérieur.*

*Une grande partie, probablement l'essentiel, de l'information qui leur serait réellement utile est celle qui concerne cet environnement extérieur et ne peut provenir que de sources extérieures" [Simon, 1990, p. 120]*

Deux types de sources sont distingués : les publications et les compilations statistiques se rapportant à l'économie. A ces deux types de sources correspondent approximativement deux sortes d'utilisation de l'information : le renseignement et la planification stratégique.

**L'information de renseignement est essentiellement utilisée pour attirer l'attention et mesurer les paramètres.** Ces systèmes de renseignement sont lents à se développer parce que :

- 1 : l'essentiel de l'information que l'on peut vouloir réunir dans un système de renseignement n'est disponible que **sous la forme de texte**.

- 2 la difficulté technique du filtrage et de la diffusion sélective de l'information supposent la capacité de "comprendre le texte". A celle-ci s'ajoute celle de rendre disponibles, dans un système informatique, des sources d'information telles que la correspondance et la conversation.

Les bases de l'articulation entre gestion et documentation, dont nous constatons l'opposition au début de cette étude, sont ici posées au niveau de la définition de l'activité de renseignement.

### c) Critique de la planification comme prévision

Il s'agit d'attirer l'attention et de mesurer des paramètres afin de parvenir à une compréhension du fonctionnement du système. Les modèles de prévision prennent une importance démesurée :

*"Prévoir pour prévoir est un jeu coûteux et inutile. L'intérêt de la prévision et des projections dans le futur devrait être de fournir une base aux décisions qu'il faut prendre aujourd'hui.*

*(...) Le futur n'intervient dans les décisions d'aujourd'hui que dans la mesure où ces décisions ont sur lui des conséquences en un certain sens irréversibles, qu'on ne peut défaire." [Simon, 1990, p.123]*

C'est la raison pour laquelle la prévision est si étroitement liée aux décisions concernant les structures physiques. En revanche, prévoir les conséquences techniques, économiques, sociales, et les progrès scientifiques est plus délicat. C'est pourtant un facteur important d'amélioration de la qualité des prévisions. Celles-ci sont essentielles à la planification.

*"La planification stratégique tournée vers les facteurs-clés de la décision formule des exigences en matière d'information très différentes de celles de la planification prise en tant que "prévision complète" et contrôle d'un système. Dans le premier cas il importe avant tout de comprendre l'architecture du système plutôt que de disposer de masses de données portant sur l'état actuel du système". [Simon, 1990, p.124]*

L'analyse de Simon définit le rôle de l'activité de renseignement dans la prise de décision stratégique alors que l'analyse précédente (l'analyse concurrentielle) faisait du renseignement un simple instrument au service de la stratégie. Cette activité porte bien encore sur l'étude de l'environnement mais elle l'analyse par des informations qualitatives, de forme linguistique, essentiellement textuelles alors que la planification repose sur des compilations statistiques. Elle prend un sens également en se distinguant de la planification et de la prévision : un système de renseignement orienté vers "la prévision complète" représente ce qu'il faut bannir au profit d'un système permettant "une compréhension générale".

### Conclusion

On comprend mieux l'ambiguïté que nous soulignons au commencement de cette première partie. La gestion stratégique est à l'origine de la "veille concurrentielle et commerciale". Il reste que c'est une certaine remise en question de la planification stratégique qui est à l'origine du développement de l'intelligence économique et d'une surveillance de l'environnement basée sur une certaine conception de la stratégie.

## **I-2-1.2 La contestation de la planification stratégique : discussion sur le rôle et la nature de l'information dans la décision**

Plus loin que la critique que fait Simon de la planification comme "prévision complète", c'est l'idée même de planification stratégique qui est remise en cause par certains. L'intelligence économique serait une nouvelle attitude vis-à-vis du futur issue d'une double critique de la prévision et de la prospective : la projection dans le futur se base sur une surveillance des événements et le modèle de l'alerte basé sur un traitement de l'information disponible remplace celui des scénarios possibles ou probables.

Les arguments de cette remise en cause viennent au moins de deux origines :

- l'une émanant de la gestion elle-même
- l'autre venant de la sociologie. Ce courant n'a pas les méthodes de planification pour objet d'étude. Il rappelle les limites d'une rationalisation de la prise de décision, non plus du point de vue cognitif (l'idée de rationalité de Simon) mais du point de vue de la dimension sociale du renseignement et de la décision stratégique.

### a) Le point de vue de la gestion : "Les hypothèses de base cachées derrière la planification stratégique" [Mintzberg, 1994, p. 232]

Il est intéressant de noter l'analyse que fait Mintzberg des "erreurs fondamentales de la planification stratégique" en relevant, en particulier, "*ses hypothèses de base cachées*". On peut y voir, en effet, autant d'arguments militant pour le développement de l'intelligence économique.

Quatre hypothèses de base sont sous-jacentes à la planification selon l'auteur :

- l'hypothèse de formalisation selon laquelle le processus stratégique peut être programmé par l'intermédiaire de systèmes afin d'éviter notamment l'initiative individuelle.

Mintzberg rapporte les propos du responsable de l'une des plus grandes entreprises américaines (Bell & Howell), qui prétend que "*c'est à la planification stratégique d'assurer que toute l'organisation connaît bien les exigences des consommateurs, les directions dans lesquelles les besoins et les attentes des consommateurs évoluent, la façon dont la technologie*

avance, et la manière dont les concurrents servent leurs clients" (propos de Marquardt [1990]<sup>23</sup> rapporté par Mintzberg).

Ce formalisme du plan s'oppose à une conception de la stratégie adaptable aux circonstances et pour laquelle existe un système d'information *ad hoc*, surveillant systématiquement l'environnement conçu pour alerter les décideurs et orienter leur décision. Bref, c'est précisément à cet archétype de la stratégie basée sur le plan que s'est construite l'idée de "competitive intelligence system".

- l'hypothèse du détachement entre la pensée et l'action : elle se manifeste par l'opposition de la gestion stratégique et la gestion opérationnelle (*ibid*, p.265). Ce modèle s'oppose à celui de l'intelligence économique qui définit un système d'information orienté pour la décision et l'action (c'est le sens du mot "*intelligence*") stratégique.
- L'hypothèse de quantification (voir aussi p.270) selon laquelle les données quantitatives auraient un privilège sur les données qualitatives. Selon Mintzberg: "*Ce message était évident dans les premières années de la littérature sur la planification, qui mettait l'accent sur la prévision et les analyses quantitatives des coûts et des bénéfices; il est toujours évident aujourd'hui, avec l'intérêt que suscite actuellement l'analyse concurrentielle et l'analyse de la valeur de l'entreprise pour les actionnaires (qui suppose l'existence de relations mesurables entre les stratégies et les prix des actions).*

*Le fait que les données soient dures signifie qu'elles sont documentées de façon non ambiguë, ce qui veut en général dire qu'elles ont déjà été quantifiées. Les planificateurs et les managers peuvent ainsi rester assis dans leurs bureaux et être informés. Pas besoin de sortir et de rencontrer les troupes, ou les consommateurs, pour découvrir comment on achète des produits ou comment on livre des combats, ou pour trouver ce qui relie les stratégies au prix de l'action ; toutes choses qui ne font que gâcher un temps précieux. Nous sommes après tout à l'âge de l'ordinateur. Des systèmes feront le travail (...)" [*ibid*, p.268-269].*

A l'hypothèse de quantification de la planification stratégique, l'intelligence économique oppose la prise en compte de l'information qualitative, linguistique, décrite dans des documents ou bien dans des connaissances tacites.

---

<sup>23</sup> Marquardt "Strategists Confront Planning Challenges" *The Journal Of Business Strategy* mai-juin 1990 p.4-8

- l'hypothèse de la prédétermination [*ibid*, p.235] reposant sur l'idée que le contexte du processus stratégique est stable s'oppose à une conception de la stratégie adaptable et flexible, basée sur un système d'intelligence économique scrutant en permanence l'environnement pour y analyser menaces et opportunités.

### b) L'apport de la sociologie

L'étude du fonctionnement des systèmes industriels, tel les districts italiens, fondés sur l'appartenance à une filière industrielle et une localisation communes, conduit à se réinterroger sur la nature de l'information entrant dans la prise de décision stratégique et donc sur la nature des systèmes de renseignements élaborés dans cette perspective. Ce ne sont plus les méthodes de prévision et de prospective qui sont en question ici mais l'échange social et l'identité collective qui conditionnent la décision stratégique de manière plus prégnante encore que l'échange économique.

Saglio [1991] a ainsi mis en évidence, à partir d'études portant sur la région Rhône-Alpes et en particulier la zone d'Oyonnax pour laquelle nous avons nous-mêmes travaillé dans le cadre de cette thèse, que dans un tel système la stratégie des acteurs ne repose pas nécessairement sur l'échange marchand classique. Selon l'auteur, ce qui distingue l'échange marchand et l'échange social c'est que, dans l'échange social, la contre-prestation que peut attendre le prestataire initial n'est pas strictement spécifiée au moment où se noue l'échange. "*De ce fait, alors que dans l'échange économique de marché ou de contrat, les deux protagonistes peuvent se considérer comme libres l'un vis-à-vis de l'autre une fois les conditions de l'échange fixées et l'échange réalisé, dans de telles situations, l'échange social crée des liens et des engagements entre les individus et ainsi renforce les liens sociaux*". [1991, p.538]

Le partage des informations commerciales et économiques sur des marchés étrangers comme la diffusion des innovations technologiques peut se réaliser sans contrepartie immédiate tant que sont respectées les hiérarchies de compétences admises.

La soumission de l'individu au groupe le protège aussi des aléas des affaires :

*"On a souvent noté, dans diverses études de cas, qu'il était possible pour l'entrepreneur de se relancer dans les affaires, même après une expérience malheureuse. Le banquier oyonnaxien de l'entre-deux-guerres et de l'après-Libération consentait à couvrir des factures non payées sur des marchés étrangers, permettant ainsi à de petites entreprises de repartir malgré de*

*difficultés financières. Les renseignements qu'il prenait lorsqu'il permettait le démarrage d'une entreprise ou la continuation de l'activité industrielle ne portaient pas centralement sur les données économiques et techniques des projets qui lui étaient proposés. Plutôt il s'inquiétait de bien identifier les réseaux de sociabilité auxquels participait le demandeur. La garantie qu'il recherchait et l'estimation qu'il faisait de la fiabilité des projets reposaient plus sur la solidité des liens d'appartenance de l'individu au groupe des entrepreneurs locaux que sur des compétences techniques et financières individuelles<sup>24</sup>. Pour le passé, les témoignages recueillis lors de notre enquête pourraient n'être que des mythes et travestir notablement la réalité. Mais certains des actuels banquiers locaux se retrouvent tout à fait dans cette description et nous ont confirmé que telles sont bien les procédures effectives qu'ils utilisent pour faire l'évaluation des projets industriels qui leur sont présentés." [ibid, p. 540]*

## I-2-2 LA GESTION DE LA TECHNOLOGIE

Une partie des réflexions sur la gestion de la technologie ont abouti à la notion de veille technologique.

Dans ce courant, Morin [1985] part du constat de "la turbulence technologique" [p. 38] pour en arriver à l'idée de veille technologique. La turbulence technologique contemporaine se caractérise par :

- La quantité d'innovations : *"l'homme a rarement innové autant, effectué des progrès en aussi grand nombre qu'aujourd'hui"*
- le caractère transversal des technologies nouvelles : elles ne concernent pas une seule activité, un seul métier. L'invention de la canette n'a eu de retombée que sur la fabrication de la soie; la micro-informatique, au contraire, s'introduit partout et modifie les conditions de fabrication dans toutes les branches.
- leur nature combinatoire : les technologies nouvelles ne peuvent être utiles qu'associées à d'autres technologies. Le micro-processeur en l'état ne sert à rien, contrairement au marteau ou au ciseau.
- leur caractère contagieux : elles ont un effet d'entraînement et un effet multiplicateur de leurs performances. Là encore l'exemple de la micro-informatique sert d'exemple dans la manière dont elle a modifié la machine-outil.

---

<sup>24</sup> C'est nous qui soulignons les propos de l'auteur en mettant en gras.

Cette spécificité des technologies nouvelles impose un "véritable défi technologique" qui est "moins dans la capacité des entreprises à inventer que dans la capacité de leur management à détecter les menaces en temps voulu, les transformer en opportunités innovatrices pour leur propre développement, à adapter les comportements individuels ou collectifs au changement qu'elles imposent." [Morin, 1985, p.39]

Pour répondre à ce défi un système interne d'informations technologiques, à l'image des "Business Intelligence Systems" des américains doit être mis en place.

La veille technologique est un outil mis au service d'une réflexion stratégique qui a le souci de définir les facteurs critiques à surveiller. Il va au-delà de ce que fait traditionnellement la documentation. Son objectif est de détecter le moindre signal d'alerte d'un changement de l'environnement.

**Tableau 15 : La veille technologique selon Morin [1985, p. 141]**

Un exemple de collecte d'informations. **Quelles informations collecter ?**

Ensemble, les responsables du marketing, de la production, de la recherche, définissent dans le cadre de la réflexion stratégique de l'entreprise :

**- les technologies à surveiller**

**- les domaines de surveillance**

Exemple: Fabrications d'aimants permanents :

<b>Domaines</b>	<b>SCIENTIFIQUE ET TECHNIQUE</b>	<b>AFFAIRES</b>	<b>ENTREPRISES</b>
<b>Technologies</b>			
<b>Matières premières</b>	Matériaux magnétiques, céramiques, alliages	Nouveaux matériaux, production, marché	Fournisseurs
<b>Procédés mis en forme</b>	Broyage, mis en forme, frittage, fonderie d'alliages, aimantation	Nouveaux procédés, nouveaux matériaux	Concurrents
<b>Applications</b>	Moteur électrique, haut-parleur, magnétron, radar	Nouveaux produits marchés	Clients

Nous retrouvons ici la même idée que celle développée précédemment à propos de la gestion stratégique. La veille technologique remplacerait alors la prévision : "pour gérer l'incertitude (...) préférer la vigilance à la prévision", affirme Morin [1885, p. 129]

b) Le mouvement de la gestion de la technologie reste orienté vers l'information sur l'extérieur (comme dans l'approche de la stratégie par les forces concurrentielles). La notion

de technologie n'est pas pensée en termes de ressources informationnelles propres à capitaliser. L'information reste centrée sur l'environnement.

## CONCLUSION

Nous avons constaté au début de notre recherche l'apparente confusion qui entoure les notions de veille et d'intelligence économique alors que ces problèmes sont des objets d'étude d'un intérêt croissant, qui suscitent une véritable explosion du nombre de publications, de formations et de manifestations, et ceci particulièrement en matière d'intelligence économique.

Au terme de la partie I-1 nous avons constaté que ces questions qui semblent à la pointe de l'actualité datent d'une trentaine d'années. Nous avons montré en quels termes elles étaient décrites. Nous avons ensuite mis en évidence la cohérence de ces problèmes en constatant qu'ils prenaient racine dans une remise en question des méthodes de planification et de prospective à la fois au niveau des entreprises et des Etats. Nous en avons conclu que l'intelligence économique est une réponse à cette remise en question.

Notre démarche a également permis de pointer les différences et les complémentarités existant entre la gestion stratégique et la gestion de la technologie sur ce point.

### ***I-3 ROLE DE LA "FONCTION DOCUMENTATION" ET DES SYSTEMES DOCUMENTAIRES***

#### **I-3-1 DOCUMENTATION ET INTELLIGENCE ECONOMIQUE**

Les systèmes de veille stratégique élaborés par les gestionnaires (par exemple ceux issus des travaux de Lesca en France) ont l'inconvénient de s'enfermer dans un cercle : analyser l'environnement en fonction d'une vision stratégique indiquant à la fois quelles sont les données à recueillir et comment les utiliser. Cette vision stratégique suppose également un consensus autour de ce qui peut être considéré comme une menace ou une opportunité que le système d'information mis en place est censé détecter.

Le rôle des systèmes documentaires est plus modeste et pourtant très étendu. Il est plus modeste dans la mesure où son processus de traitement de l'information ne dépend d'aucune vision stratégique mais seulement de l'interprétation que les utilisateurs en feront. Il est très

étendu en ce sens que l'essentiel des connaissances produites et échangées fait ou est susceptible de faire l'objet d'un document. L'exploitation systématique et non empirique de ces gisements de documents produits en interne ou disponibles à l'extérieur est donc précieuse pour un système de renseignement destiné à la décision stratégique, bref pour l'intelligence économique.

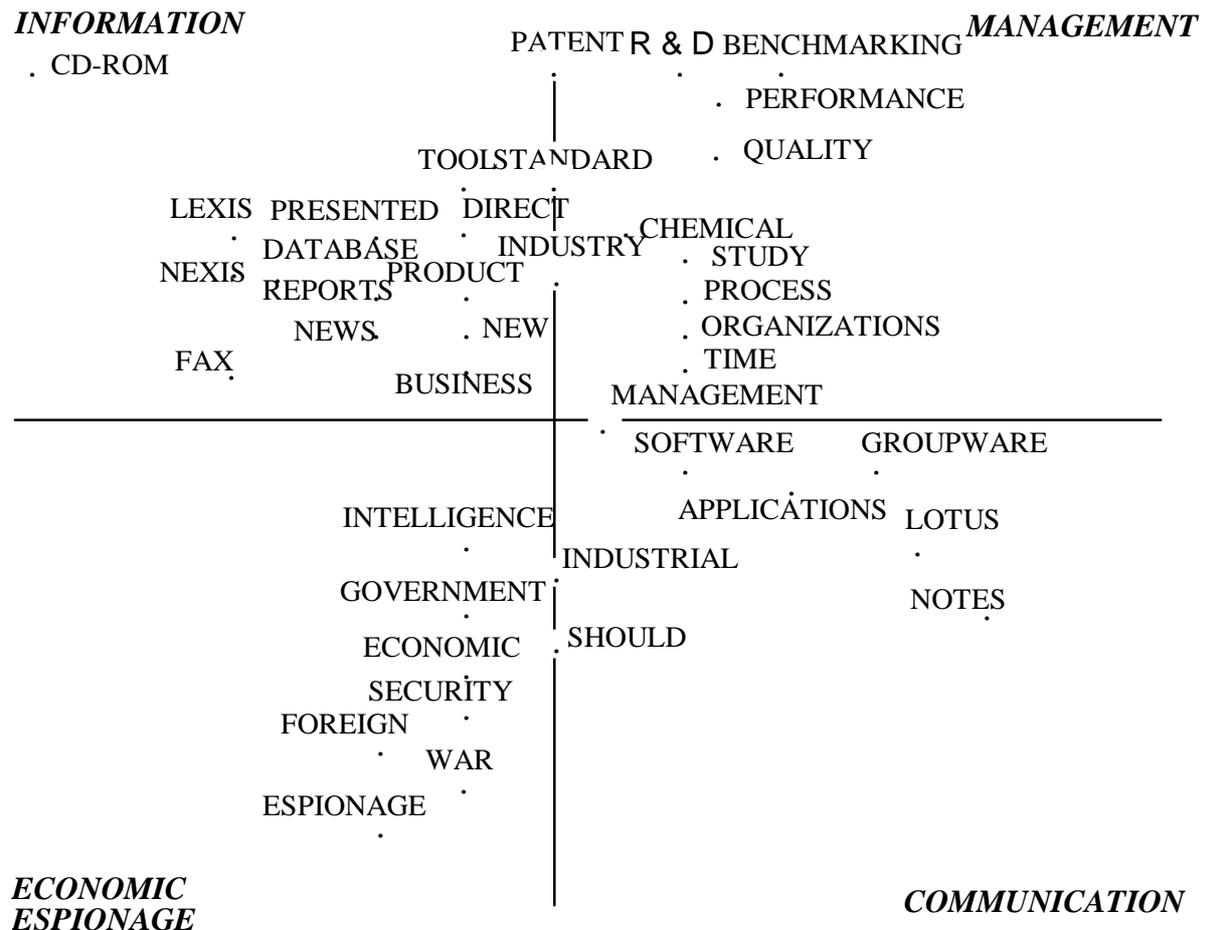
Sollicitée en ce sens, la fonction "documentation" devrait prendre une ampleur particulière dès lors qu'elle n'est plus seulement liée à la conservation des documents et à la valorisation d'un fonds.

L'intelligence économique semble davantage intégrer le rôle de la documentation que ne l'a fait la veille stratégique. Besson et Poussin [1996] lui attribuent la fonction "mémoire" de l'entreprise, Bloch [1996] rappelle la fonction de la bibliométrie pour améliorer les performances de la documentation dans le cadre de l'intelligence économique. Les publications américaines comme celle de Fuld dans *"The New Competitor intelligence"*[1995] sont des répertoires de sources d'informations. On peut dire que la documentation et les services commerciaux sont donc les premiers centres d'information sur le monde extérieur à l'entreprise. Les promoteurs des réseaux du type "groupware" ont beaucoup insisté sur cet aspect, au point d'en faire, lors d'une étude que nous avons faite en 1995, l'un des vecteurs privilégiés de l'intelligence économique. La mise en commun des ressources documentaires et de l'information des commerciaux au moins pour le partage des renseignements sur les clients si ce n'est pour celle de la connaissance de la concurrence au service de la stratégie de l'entreprise.

### I-3-2 ETUDE DE 1995 SUR LE THEME "COMPETITIVE INTELLIGENCE"

Cette étude réalisée à partir des bases de données américaines décrites précédemment (I-1-1) nous avait conduit à réaliser une analyse factorielle des résumés des articles contenant la notion "competitive intelligence". Les résultats du tableau 16 résument une partie de cette analyse.

**Tableau 16 :**  
**Analyse factorielle des correspondances : *Intelligence économique 1995***  
**Carte des facteurs 2 et 3.**



On voit la notion d'intelligence économique au centre de quatre grands thèmes :

- les produits d'information
- les outils de travail en groupe et de partage de l'information (groupware)
- le benchmarking, traduit en général par "étalonnage concurrentiel" et qui consiste à étudier la stratégie de ses concurrents les plus performants pour les imiter : l'intelligence économique est l'outil de cette étude.
- l'espionnage économique : il s'agit de montrer quelles sont les limites de ce qui fait l'objet de l'intelligence économique. Si cette dernière est concernée par la collecte et le traitement

de renseignements sur la concurrence ainsi que par les stratégies défensives (la "contre-intelligence"), elle ne doit pas néanmoins être confondue avec l'espionnage ("economic espionage") qui poursuit des objectifs similaires mais repose sur des méthodes illicites et "non éthiques".

Produits et systèmes documentaires d'une part, gestion d'autre part (benchmarking) semblent bien articulés dans cette conception américaine de l'intelligence économique.

### I-3-3 CONSTAT SUR L'ORGANISATION DES STRUCTURES DE DOCUMENTATION EN FRANCE :

L'implantation de la veille comme de l'intelligence économique dans les entreprises s'effectue, d'après ce que nous avons pu constater sur le terrain, dans les centres de documentation : même si les services d'intelligence économique sont parfois nés dans les services marketing (services réalisant les études de marché) ils finissent par s'organiser autour des centres de documentation.

Toutefois l'ensemble des fonctions de l'entreprise collectant et produisant de façon systématique de l'information sur l'environnement ne sont pas, à notre connaissance, coordonnées. Les services d'analyse de la concurrence, de documentation, de veille scientifique, d'information économique, de recherche brevets sont "atomisés" c'est-à-dire conçus comme des centres répondant chacun à la commande d'une personne autorisée mais sans relation les uns aux autres. L'absence de coordination et de réseau entre les différents services dont la mission est d'informer nous a semblé important.

### I-3-4 EVOLUTION DE LA DOCUMENTATION

L'intelligence économique exige une évolution des performances attendues de la documentation. La documentation n'est plus seulement pensée comme un centre de ressources (de livres, d'accès à des bases de données et CD-ROM) mais comme une fonction par laquelle

des connaissances sont capitalisées, extraites des documents et retrouvées. On parle aujourd'hui de "gestion des connaissances".

L'évolution des idées concernant le rôle de la documentation semble nous mener de l'extérieur vers l'intérieur. L'information stratégique n'est plus réduite à l'information sur l'environnement mais elle est redéfinie par rapport aux ressources internes en matière de connaissances et de savoir-faire dont dispose une entreprise<sup>25</sup>. Or ces ressources internes ne se limitent plus à "l'information informelle", celle qui n'est publiée sous aucune forme. L'accent est mis sur la formalisation des savoirs internes. Après s'être concentrée sur l'approche de la stratégie par les forces concurrentielles on semble se focaliser sur l'importance des savoirs et des méthodes nécessaires à leur capitalisation. Or cette capitalisation se matérialise dans des documents.

#### ***I-4 CONCLUSION***

Nous avons pu constater au cours de cette étude que l'intelligence économique n'était pas *stricto sensu* une "nouvelle mode" mais un concept trouvant son origine dans les théories américaines du management à la fin des années 1960. Pourtant c'est au début des années 1990 qu'il se développe en France. Il ne s'agit en aucun cas, nous semble-t-il, d'un retard de la France dans la prise de conscience du rôle de l'information dans l'économie. Les rapports officiels et les politiques n'ont pas manqué sur ce sujet en France et, cela dès le début des années 1960 [Salmon 1963]. Depuis le rapport Salmon jusqu'au rapport Martre [Plan 1994] en passant par les rapports Lenoir-Prot [1979], Meyer [1990] etc., qu'est-ce que l'introduction de la notion d'*intelligence économique* a apporté de nouveau ?

Il semble, comme nous l'avons déjà mentionné, que ce soit la dimension offensive liée à la gestion de l'information qui motive l'appel à l'intelligence économique. L'intelligence économique est une forme de guerre, celle que les Etats-Unis dominent grâce à la place prépondérante qu'ils occupent dans le secteur de l'information et que l'information occupe

---

<sup>25</sup> On peut consulter à ce sujet par exemple : Charue-Duboc (sous la direction de)[1995]; Revue Française de Gestion 1995 n°105.

dans l'économie. Notre engouement pour l'intelligence économique dans les années 1990 est "réactif" par rapport à cette situation.

D'après Schiller [Le Monde Diplomatique du 20/08/1997] "*on estime à plus de 1000 milliards de dollars depuis 1945 les aides [du gouvernement fédéral] qui ont permis entre autres, le rapide développement des ordinateurs et des secteurs de l'informatique et de l'intelligence artificielle. Ces industries, ainsi que les domaines qui s'y rattachent, ont largement contribué à la supériorité américaine dans les technologies de l'information, les réseaux informatiques, la création de bases de données, l'industrie des effets spéciaux et des systèmes de surveillance internationale, c'est-à-dire l'infrastructure de ce que l'on nomme aujourd'hui "l'ère de l'information"*" [1997, p.20-21]. Le résultat serait que sur les 1000 milliards de dollars de chiffre d'affaires mondial que représente l'informatique aujourd'hui, 80% serait détenu par les USA [Le Monde du 10/08/1998]

Aux Etats-Unis c'est au sein des sciences de gestion que les notions de "competitive intelligence" et "environmental scanning" sont nées. Une réflexion sur le rôle de l'information pour les acteurs économiques, son intégration dans la gestion des entreprises se sont donc développées *en même temps* que les technologies de l'information.

## **ANNEXE I-A**

### **Liste des attentes ("*want-list*") en matière d'information relevant de l'intelligence économique selon Cleland et King [1975]**

#### *- Marketing Information :*

- Pricing, discounts, terms and product specifications
- Volume, history, trend and outlook for a given product
- Market share and trend
- Marketing policies and plans
- Relations with customers
- Size and deployment of sales force
- Channels, policies and methods of distribution
- Advertising program

#### *Production and Product Information*

- Evaluation of quality and performance
  - Breadth of line
  - Processing and technology
  - Product cost
  - Production capacity
  - Location and size of production facilities and warehouses
  - Packaging Delivery Research and development capability

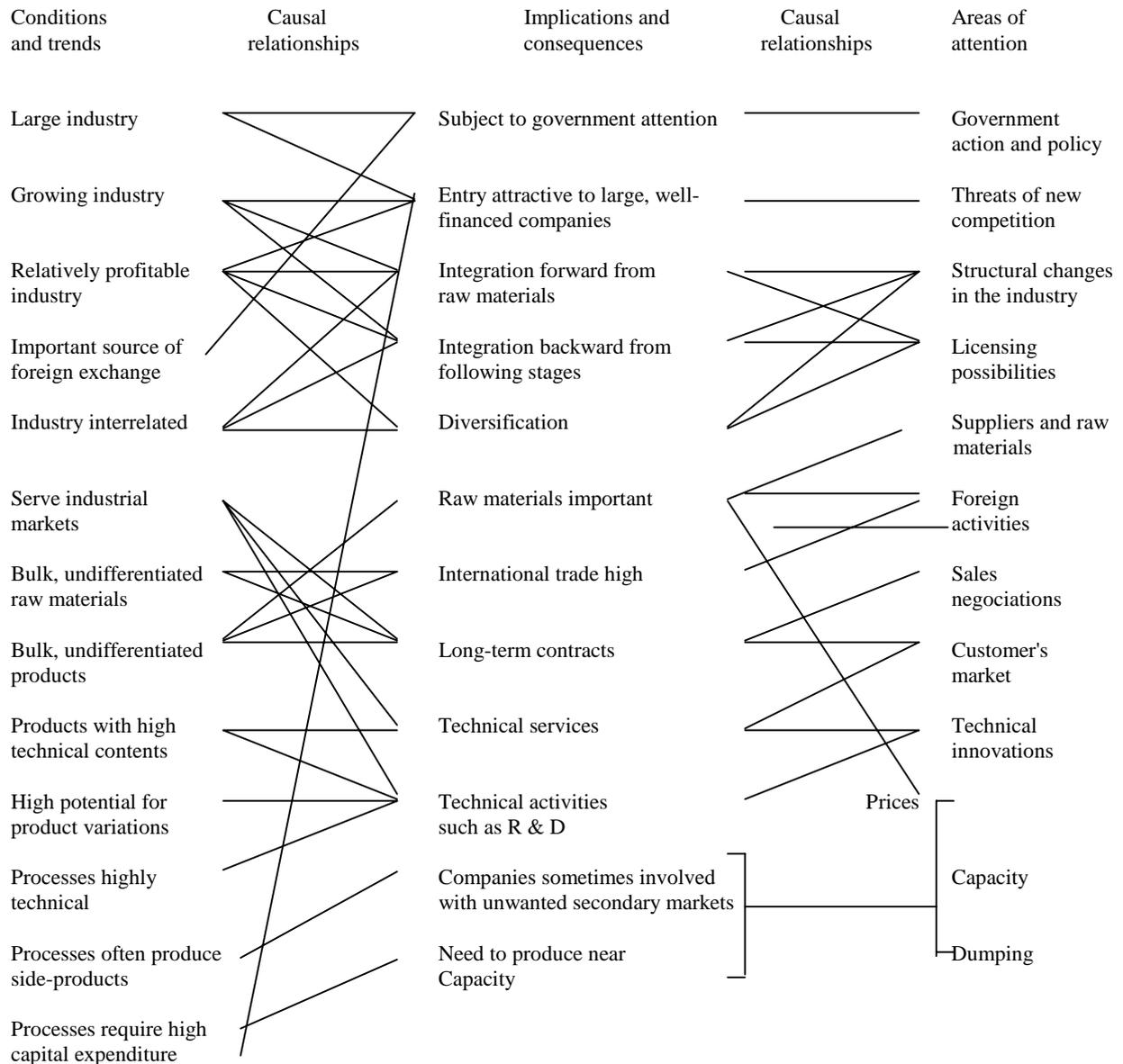
#### *Organizational and Financial*

- Identification of key decision makers
- Philosophies of key decision makers
- Financial condition and outlook
- Expansion and acquisition programs
- Major problems/opportunities
- Research and development programs

## ANNEXE I-B :

### The relationships of Industry Conditions to Areas of Scanning Attention

(basé sur plusieurs secteurs de l'industrie chimique)



\*Based on several sectors of the chemical industry.

Tableau d'après Aguilar [1967]

**Partie II**  
**LE TRAITEMENT DE L'INFORMATION TEXTUELLE,**  
**OUTIL DE L'INTELLIGENCE ECONOMIQUE**

## PARTIE II : LE TRAITEMENT DE L'INFORMATION TEXTUELLE, OUTIL DE L'INTELLIGENCE ECONOMIQUE

### II-1 LE PROBLEME

#### II-1 -1 QU'EST-CE QU'UN DOCUMENT ?

Un document est, au sens étymologique (issu du latin *documentum*), "ce qui sert à instruire": c'est le sens juridique de "document" entendu comme "écrit servant de preuve" qui en dérive directement. Cela signifie qu'est constitué comme document tout ce qui est susceptible de déceler une information et mérite, à ce titre d'être conservé et stocké. La valeur de vérité de ce qui est contenu dans un document n'est pas en question. Quelque chose est *constitué* comme document : il y a là un traitement par lequel cette opération de constitution est effectuée.

Le traitement du document fait donc partie du concept même de document. Il consiste à :

- à établir une "fiche d'identité" d'un contenu et de sa source (traitement documentaire par lequel un objet est transformé en document : cette opération s'appelle dans le milieu des bibliothèques "le catalogage")
- à les stocker sur un support en vue de leur conservation (traitement physique)
- à les insérer dans une organisation visant à les retrouver (traitement de l'information).

Dans le cadre des traitements informatisés, il résulte de ceux-ci une double structuration du document :

- une structuration logique : les formats définissant le type de document (traitement de texte, base de données, image, tableur de tel type, HTML etc.). Il y a même une définition informatique du document qui ramène celui-ci à un fichier associé à un format particulier.
- une structuration associée à la description du document et de son contenu : c'est la fonction des normes de catalogage et des métadonnées. Cette structuration découpe le document en champs : numéro, titre, auteurs, source, année, localisation géographique du laboratoire ou des auteurs, l'entreprise concernée, descripteurs, résumé, etc.. Ces champs forment autant d'informations "secondaires" par rapport au contenu (le résumé ou le texte intégral quand il s'agit de document textuel). Le champ "mots-clés" a consisté à créer une sorte d'intermédiaire entre le texte et les données secondaires en ramenant celui-ci à des

événements indépendants décrivant le contenu du texte. La normalisation de cet ensemble d'informations secondaires joue aujourd'hui un rôle capital avec le problème du catalogage des ressources d'Internet.

Grâce à cette double structuration les documents peuvent être échangés et retrouvés.

Pour être retrouvé le document doit être indexé. L'indexation, rendue possible par la double structuration, consiste à constituer pour chaque élément du document une adresse permettant, à partir d'une requête, de le retrouver.

La requête est la représentation de la question que l'utilisateur pose au système documentaire. Elle est une représentation en ce sens que la question ne pourra pas toujours être formulée "en langage naturel" (c'est-à-dire telle qu'il la formulerait avec un interlocuteur humain) mais devra respecter des contraintes pour répondre à un modèle "compréhensible" par la machine (par exemple le modèle booléen auquel s'ajouteront tel ou tel type de possibilité de troncage, de marqueurs de proximité entre les termes etc.).

Selon la structuration logique choisie une série de documents peut être une base de données (voir les logiciels documentaires comme Texto, Alexandrie etc.), une base de données relationnelle (par exemple un système comme Dbase) un fichier unique ou une série de fichiers. L'objectif est de pouvoir non seulement retrouver l'information mais aussi la chaîner (retrouver tous les documents dont la société est  $x$  et dont la source est  $y$ ) plus ou moins "intelligemment" c'est-à-dire en offrant plus ou moins de possibilités de chaîner, non pas seulement des données secondaires ou des chaînes de caractères, mais aussi des contenus élaborés.

Ce chaînage des informations particularise la recherche documentaire informatisée.

Le document électronique se distingue en outre du document imprimé par :

- Son mode de lecture : l'hypertexte. La notion même de document devient difficile à être sauvegardée puisque non seulement les documents peuvent renvoyer les uns aux autres mais le document est découpé en éléments d'information liés à d'autres éléments de documents étrangers.
- Le multimedia c'est-à-dire la possibilité d'ajouter le son, l'image animée et à plusieurs dimensions au texte et à l'image en deux dimensions
- La production et la consultation en réseau.

## II-1-2 LA VALEUR STRATEGIQUE DES DOCUMENTS

Un document est donc un objet complexe (il comprend des informations primaires et secondaires auxquelles correspond une structure). Il n'est pas assimilable à une donnée. Une donnée est brute, ne contient rien en elle-même et ne fait pas sens en elle-même. Le document est complexe car il contient une connaissance, c'est-à-dire un ensemble de données élaborées qui fait sens par lui-même.

La prise en compte de la valeur stratégique de l'information porte souvent sur celle des données [voir par exemple Drèze 1995] : les données issues des scanners dans le secteur de la grande distribution ou encore l'analyse des chèques, panels, enquêtes mais aussi les mesures d'audience dans l'audiovisuel etc., les mesures destinées au contrôle statistique des procédés sont celles auxquelles toute une série de méthodes quantitatives et d'outils informatiques s'intéressent afin d'en extraire des connaissances et de fournir une aide à la décision pour des stratégies d'action futures.

Les documents font partie de ces gisements d'informations dont le traitement peut avoir une valeur stratégique. Leur spécificité est de contenir des connaissances.

La critique adressée à l'idée que le traitement des documents peut avoir une valeur stratégique consiste à souligner que le document est par nature rétrospectif : il peut donc difficilement aider à une prise de décision future.

L'idée que le traitement du contenu des documents écrits, à une échelle importante, peut être un outil stratégique repose sur l'hypothèse que la combinaison d'informations par définition anciennes (et l'on peut considérer comme ancien tout ce qui ne repose pas sur un accès en temps réel) crée des informations nouvelles qui, en tant que telles sont susceptibles d'éclairer une décision. Il est trivial en effet de rappeler que tout modèle s'élabore à partir d'éléments observés donc passés. Dans notre cas il ne s'agira pas d'élaborer un modèle mais de situer le présent en fonction de l'analyse des documents du passé pour alimenter une décision sur l'avenir.

L'argument consistant à dévaloriser toute démarche de ce type sous prétexte qu'une information publiée est déjà périmée est en contradiction avec les idées mêmes de veille et d'intelligence économique telles que nous les avons déjà définies dans la première partie de cette étude.

C'est précisément pour répondre à une crise des modèles, méthodes et organismes de prospective qu'ont été mis en place des dispositifs de veille et d'intelligence économique à savoir des dispositifs reposant strictement sur la collecte et le traitement systématique des informations disponibles. L'alerte (la capacité d'alerter l'organisation en cas de changement concernant des opportunités ou menaces) a pour vocation de remplacer les modèles de prévision en matière stratégique. Que l'objectif soit ou non atteint est une autre question (celle de l'évaluation des dispositifs de veille et d'intelligence économique) mais leur raison d'être est celle que nous venons d'évoquer. On ne peut donc arguer contre ces dispositifs qu'ils utilisent seulement de l'information, qui en elle-même n'aurait pas de valeur sans modèle de prévision !

Toutefois chaque type de document s'inscrit dans une temporalité qui lui est propre et qui doit être connue et maîtrisée pour donner aux résultats de leur traitement tout leur sens. L'information financière est accessible en temps réel. L'information économique est mise à jour hebdomadairement. L'information scientifique est mise à jour par trimestre, voire par semestre, et repose pour la partie "publications" sur des informations qui ont elles-mêmes suivi un délai de diffusion, avant publication, de un à deux ans pour les revues à comité de lecture internationaux. Les brevets sont publiés après un délai de 18 mois à partir de la date de dépôt.

Les documents dont nous envisageons ici le traitement sont des documents textuels.

## **II-2 LES TRAITEMENTS POSSIBLES DU DOCUMENT TEXTUEL**

On peut différencier trois approches possibles :

- recherche documentaire
- bibliométrie
- analyse de contenu

## II-2-1 LA RECHERCHE DOCUMENTAIRE

La recherche documentaire a pour objectif la sélection des documents en fonction d'une requête ou de son paraphrasage. La méthode consiste à aider à la reformulation de la requête en proposant à l'utilisateur des termes supplémentaires et/ou un traitement en langage naturel de sa requête jusqu'à rejoindre les mots décrivant les documents. Une telle approche est donc centrée sur le traitement automatisé du langage (T.A.L) puisqu'elle est centrée sur la requête. Une approche plus récente a consisté à déplacer le centre d'intérêt de la requête à l'usager (modélisation des utilisateurs) avec la même idée qui consiste à aller de l'usager, ou de sa requête, aux documents.

## II-2-2 L'ANALYSE DE CONTENU

Ce que nous appelons "analyse de contenu" relève d'un paradigme que nous pourrions dire inverse. Elle a pour objectif la réponse à un besoin d'information en fonction du contenu des documents et non en fonction de la requête. Ce sont les documents et non la requête qui sont analysés d'où le privilège des méthodes d'analyse de données dans cette perspective. La requête joue le rôle de sélecteur d'une sous-base (voir III-1-1 p.118), celle contenant toutes les réponses possibles au sujet. L'usager doit situer sa requête en fonction du contenu des documents.

La distinction entre recherche documentaire et analyse de contenu peut paraître délicate. En effet, l'indexation (entendue comme la sélection des unités textuelles qui représentent le contenu du document) et la recherche d'information (appariement entre termes d'indexation du document et termes de la requête) peut appartenir aux deux domaines : il s'agit d'apparier des contenus, celui de la requête avec ceux des documents.

En réalité la différence est, selon nous, dans la méthodologie choisie. La recherche documentaire relève le plus souvent de l'informatique et de l'ingénierie linguistique. Comme telle, son approche est CENTREE SUR LA REQUETE et plus largement sur L'USAGER tandis que l'analyse de contenu est CENTREE SUR LES DOCUMENTS. En effet, l'informatique documentaire recherche la satisfaction de l'usager du système en tentant d'enrichir les termes de la requête pour qu'ils puissent rejoindre ceux employés dans les documents : l'informatique documentaire se concentre donc sur l'aide à la reformulation,

l'expansion automatique de la requête, la constitution de thesaurus, la modélisation des utilisateurs pour pré-orienter la recherche, le feed-back ou l'interactivité entre l'utilisateur et le système etc. Bref, on cherche à organiser l'appariement entre requête et documents à partir des termes de celle-ci et même plus encore à partir du besoin d'information de l'utilisateur voire d'un modèle de l'utilisateur lui-même.

L'analyse de contenu, telle que nous l'envisageons, fait exactement l'inverse : elle part du contenu de la base de documents et elle impose ce contenu au demandeur d'information en exigeant qu'il se positionne lui-même par rapport au contenu existant. Cette démarche s'adresse davantage aux experts car ils ont déjà la maîtrise du champ dans lequel leur besoin d'information se situe. Les experts ne cherchent pas seulement à "trouver" des documents sur un sujet mais plutôt à naviguer dans un champ de connaissances qui leur est familier.

Telle est l'approche qui sera la nôtre dans la troisième partie de cette étude. L'intelligence économique est donc pour nous un cas particulier de la recherche d'informations en milieu expert.

### II-2-3 LA BIBLIOMETRIE

Ses méthodes n'ont pas été conçues pour la documentation (la recherche d'information) mais sont utilisées par elle. L'exemple de Live Topics mis en service sur Internet avec le moteur de recherche Altavista en est l'exemple même. Il joue le rôle d'une aide à la reformulation en réutilisant <sup>26</sup> la classification Leximappe qui avait été conçue par le Centre de Sociologie de l'Innovation de l'Ecole des Mines de Paris pour cartographier le contenu d'un domaine scientifique.

Cette méthode ne recherche pas d'association entre mots et documents : elle n'a pas été pensée pour cela. Seuls les descripteurs sont examinés en tant que matériaux pour la construction d'une image de la science.

Le point commun entre documentation et sociologie des sciences, entre bibliométrie et scientométrie est l'utilisation des publications. Mais au-delà d'un objet commun partagé il y a des problèmes communs. On voit bien en effet que l'aide à la reformulation va au-delà d'un problème technique, traditionnellement celui de la polysémie (ambiguïtés) et celui de la synonymie (expansion des termes de la requête). L'analyse d'un besoin d'information à travers l'analyse de ses réponses possibles (ce qui relève de la documentation) suppose celle de l'état

de la science ou de la technique ou du marché c'est-à-dire suppose une analyse à la fois sociologique et épistémologique du besoin d'information comme du contenu des documents supposés y répondre. La documentation tend à repérer, classer, cartographier des connaissances au-delà de leur incarnation dans des documents. La veille et l'intelligence économique incarnent cette complexification des performances attendues de la documentation. On comprend donc qu'un certain partage des outils entre sociologue et chercheur en sciences de l'information ait lieu au-delà du partage des compétences techniques entre linguistes, cognitivistes et informaticiens pour améliorer les systèmes documentaires par l'ingénierie linguistique.

## **II-3 L'INFORMATIQUE DOCUMENTAIRE :**

### **II-3-1 LES METHODES**

Elles s'organisent autour de trois grandes rubriques qui peuvent se superposer :

1. les interfaces

2. le traitement de la langue naturelle

3. les réseaux sémantiques

1- Les interfaces, que l'on a cherché à améliorer essentiellement par l'hypertexte ou le graphisme pour ce qui est de la recherche documentaire, réorganisent les informations pour permettre à l'utilisateur de naviguer à l'intérieur d'un ensemble de documents avec le moins d'apprentissage possible. (voir par exemple : AGOSTI et SMEATON 1996).

2- Le traitement de la langue naturelle propose :

- une indexation linguistique qui permet de ne pas réduire les termes de l'indexation aux formes graphiques mais à des unités linguistiques complexes (le syntagme nominal <sup>27</sup> en particulier) associées à des modèles de recherche d'information permettant un appariement de ces unités.
- le paraphrasage qui consiste à déduire, à partir des termes de la requête des liens de synonymie.

---

<sup>26</sup> Source : séminaire de l'ADEST du 09/12/1997.

<sup>27</sup> Voir en particulier Bouché [1989] ; Bouché et Germain [1991] ; S.Lainé [1982].

- la reformulation et l'expansion de la requête <sup>28</sup> qui consiste à proposer à l'utilisateur d'autres termes synonymes calculés, soit à partir des documents, soit à partir d'un dictionnaire pré-constitué, mais dans une démarche pouvant comprendre une phase d'apprentissage. L'utilisateur juge l'expansion de termes proposée par le système qui "mémorise" ce jugement pour calculer d'autres relations.
  - L'approche terminologique : cette approche est utilisée en veille technologique. Elle consiste à construire des bases de connaissances terminologiques dans des domaines particuliers relatifs au domaine de l'entreprise et à rechercher des documents en fonction des arbres terminologiques constitués à partir de la requête du demandeur de l'étude.
- 3- Les moteurs sémantiques (par exemple le système *Taiga* du Centre Scientifique d'Etudes de Défense ou la dernière version de *Spirit*) : il s'agit de déduire d'un énoncé des énoncés proches même s'ils n'en contiennent aucun des mots. *Taiga* se propose de transformer les textes de n'importe quelle langue dans un langage pivot qui regroupe les terminologies autour de champs sémantiques. *Spirit*, essentiellement axé sur le traitement de la langue naturelle permettrait également dans sa nouvelle version un traitement sémantique. *Taiga* (Traitement Automatisé de l'Information Géopolitique d'Actualité) est un logiciel mis au point par Christian Krumeich pour les besoins de la DGSE et *Spirit* a été développé par le Commissariat à l'Energie Atomique.

De nombreux travaux ont été réalisés pour construire les champs sémantiques de termes à partir de leur utilisation dans une base de données<sup>29</sup>. Ainsi calcule-t-on :

- des associations entre termes, non pas à partir de cooccurrences (mots apparaissant dans la même unité de contexte) mais de relations substance/prédicat (voir par exemple [Ruge et Schwarz 1991], [Schwarz 1990]), ou de relation sujet/objet [Hirshman, Grishman, Sager 1975] afin de grouper des noms qui apparaissent comme sujet ou objet des mêmes verbes. On construit ainsi la sémantique d'un domaine spécialisé.
- Des thesaurus liant les termes par des relations sémantiques (hyponymie, antonymie, méronymie, implication, dérivation morphologique). C'est le cas de Wordnet, réseau

---

<sup>28</sup> On peut consulter à ce sujet Beaulieu 1997 "Experiments on interfaces to support query expansion", *Journal of Documentation*, 53(1) 1997, pp.8-19.

<sup>29</sup> On peut consulter à ce sujet Favier L. 1994 et Habert B. et alii 1997

lexical en anglais disponible sur le réseau Internet<sup>30</sup>. Un sens se définit par la place d'un terme dans le réseau.

- des agrégats (clusters) d'expressions [Croft et Lewis 1990] qui définissent des relations de proximité entre des expressions à l'intérieur d'une base.

### II-3-2 LES LIMITES

Les limites de la recherche documentaire résident dans le fait que, plus elle est élaborée, plus elle dépend de *prérequis* qui conditionnent la qualité des résultats : dictionnaire et thesaurus spécialisés contenant les seules ressources lexicales que le système pourra reconnaître auxquelles on ajoute une base de connaissance (un système de règles elles-mêmes limitées). L'hétérogénéité des corpus traités dans le cadre de domaines pluridisciplinaires et la spécialisation du langage en même temps que l'évolution permanente de la terminologie rendent ces approches difficiles.

De plus, les questions de l'utilisateur peuvent être plus ou moins abstraites. En effet, par définition un utilisateur ne sait pas ce qu'il peut trouver : sa question sera donc plus ou moins précise. D'autre part plus son besoin d'information est complexe, plus sa formulation comporte de concepts abstraits. Or toutes les questions stratégiques sont générales et abstraites : elles ne se réduisent pas à la demande d'un renseignement. Pour mieux le comprendre essayons d'envisager une typologie des besoins d'information.

#### *Une typologie des besoins d'information exprimés dans un centre de documentation spécialisé :*

Le travail effectué sur l'évaluation des systèmes documentaires repose toujours sur le niveau de complexité des requêtes : les usagers utilisent-ils le plus souvent un mot ou davantage ? Comment enrichissent-ils leur requête (utilisation d'un thesaurus, cheminement par essais et erreurs...) ?

---

<sup>30</sup> Par FTP anonyme depuis FTP.cogsi.princeton.edu

Mais avant de chercher à comprendre les stratégies des utilisateurs pour formuler leur requête ne peut-on pas, en amont, chercher à comprendre leur besoin d'information pour déterminer en quoi un système documentaire peut y répondre ?

L'utilisateur d'un centre de documentation spécialisé (dans une entreprise ou un centre de recherche) peut avoir des demandes d'information aussi différentes que celles que nous répertorions dans le tableau ci-dessous :

**Tableau 17 : Besoin d'information/compétences requises**

<u>besoin d'information</u>	<u>compétence requise</u>
Chercher une donnée	Connaître la source, savoir où la chercher
Chercher un brevet sur lequel le chercheur a quelques indications (il en connaît l'auteur, a une idée de l'année, du domaine et il veut en voir le contenu)	Connaître la recherche de brevet (les champs caractérisant les documents brevets, les bases de données de brevets).
Chercher des informations sur la micro-injection de polymères S'ajoute rapidement un niveau supplémentaire de difficulté : la formulation du problème. En interrogeant par "micro-injection" la base RAPRA en 1995 on obtenait 2 réponses. D'où la nécessité de faire définir par l'utilisateur le concept utilisé comme clé de l'interrogation afin de définir un réseau de synonymie. C'est pour aider l'utilisateur dans cette étape que se sont développés les travaux sur la reformulation et l'expansion de la requête. Mais ici on en reste à la recherche d'information portant sur un seul concept.	Savoir quelle est la base à interroger, connaître le langage de la base. Définir un réseau de synonymie
Faire le point sur le " <i>statistical process control</i> " dans les technologies de transformation des polymères. Là le nombre de références explose et aucun système ne peut résoudre ce problème. Limiter la question c'est l'amputer	Pas de problème de formulation, expression juste mais triviale concernant un concept important. D'où une seule possibilité : limiter arbitrairement le nombre des réponses.
Expression d'une idée formulée avec d'autres termes que ceux de la base. Rechercher " <i>textual data analysis</i> " ne permet d'obtenir aucune réponse conforme à l'idée contenue dans la question et cela bien que la traduction soit juste. pas plus que <i>factor analysis</i> pour analyse factorielle ou multidimensionnal statistics pour statistique multidimensionnelle. En revanche " <i>exploratory data analysis</i> " couplé à " <i>semantic analysis</i> " permet de trouver les réponses à la question. S'agit-il d'un problème de terminologie ? Non dans la mesure où il faut aller au-delà de la traduction des termes et expressions pour comprendre de tels problèmes. L'histoire et le contexte d'utilisation de certaines méthodes, les aspects culturels, expliquent que certains termes sont préférés à d'autres pour désigner les mêmes choses.	Connaissances culturelles. Terminologie.
Expression d'une idée contenant un concept abstrait : "les applications des polymères à cristaux liquides", le marché des polymères à cristaux liquides".	Le thème "application" étant abstrait, les documents parlant d'applications ne mentionneront pas nécessairement ce terme. D'où l'obligation d'une recherche large avec le terme "polymères à cristaux liquides" et la seule opération possible est de trier les réponses en fonction de l'idée que l'on se fait <i>a priori</i> du problème.

La recherche documentaire est parfaitement adaptée aux trois premiers cas. Peut-elle l'être pour les cas suivants ? La demande, sous forme de question, introduit d'autant plus d'*a priori*

sur la recherche qu'elle sera précise, que le nombre de mots sera grand, qu'une expansion sera riche. La démarche n'est pas aisément adaptable à une méthodologie pour l'intelligence économique car cette dernière recherche précisément la nouveauté, c'est-à-dire suppose le moins possible d'*a priori* dans la recherche.

De plus des problèmes linguistiques liés à la bonne interprétation des termes se posent : polysémie, homonymie, homographie etc.

## **II-4 STATISTIQUE SUR LES DONNEES TEXTUELLES (ANALYSE DE CONTENU) OU SUR LES DONNEES SECONDAIRES DU DOCUMENT (BIBLIOMETRIE) :**

### **II-4 -1 LA BIBLIOMETRIE, PRESENTATION GENERALE**

En 1969 Pritchard <sup>31</sup> pour les uns, Otlet <sup>32</sup> pour les autres, invente le terme de "bibliométrie" pour le substituer à celui de "bibliographie statistique" (qualifié d'ambigu car pouvant désigner une bibliographie sur les statistiques). Ce terme de "bibliographie statistique" était employé depuis les travaux de Hume 1923 <sup>33</sup>.

Selon Pritchard la bibliométrie se définit comme :

*"L'application de méthodes mathématiques et statistiques aux livres et autres media de communication".*

Pritchard ne dit rien sur la finalité de ces méthodes.

On peut toutefois distinguer trois grands courants.

- **"la bibliographie statistique"**

**Contexte** : la gestion et l'évaluation des bibliothèques.

**Objet analysé** : la circulation des ouvrages et périodiques.

---

<sup>31</sup> "Statistical Bibliography or Bibliometrics?", *Journal of Documentation*(25),1969

<sup>32</sup> D'après Estivals [1969]

<sup>33</sup> *Statistical Bibliography in Relation to the Growth of Modern Civilisation*, Edition Grafton, 44 p., 1923

La bibliométrie distributionnelle (les grandes lois : Bradford, Zipf, Lotka) est née dans ce contexte.

- **La bibliométrie des références bibliographiques**

**Contexte** : la scientométrie (l'évaluation de la recherche) et la veille technologique.

**Objet analysé** : les documents écrits (publications, brevets etc.) à travers les champs concernant l'identité des documents et non les champs textuels (d'où le terme de référence bibliographique).

- **L'analyse de contenu des documents appelée parfois "statistique textuelle"** [Lebart et Salem, 1994].

**Contextes multiples** : littéraire, sociologique, psychologique (analyse de discours), la veille stratégique (elle utilise encore peu ces méthodes et se cantonne à l'analyse des références bibliographiques le plus souvent), les systèmes documentaires (sélection des réponses dans une base de textes à partir d'une requête) etc.

**Objet analysé** : le texte (une œuvre, des réponses à des questions ouvertes, des entretiens, une base de textes scientifiques/ résumés ou textes intégraux, la presse etc.)

Le point commun de tous ces courants de la bibliométrie : la communication écrite. L'informatisation de cette communication écrite a rassemblé de nombreuses disciplines qui ont comme objet des textes ou des discours transcrits en grand nombre ; d'où la nécessité d'introduire pour leur analyse des méthodes quantitatives. Malheureusement ces disciplines s'ignorent. Même Rostaing [1996] dans un ouvrage récent présentant la bibliométrie prend partie pour le deuxième courant de la bibliométrie et affirme : "*La bibliométrie cherche à analyser les écrits scientifiques tandis que les traitements statistiques linguistiques ont été développés initialement pour l'analyse des textes littéraires*".

Il ne nous semble pas judicieux de choisir une méthodologie en fonction de la discipline à laquelle appartiennent les textes traités. C'est une direction qui est choisie par certains en

documentation par exemple (modélisation des utilisateurs passant par une modélisation des types de textes ; caractérisation des textes et recherche de normes textuelles différentes selon la discipline etc.). Or d'une part, la science est pluridisciplinaire et, d'autre part, les méthodes développées dans le cadre du traitement des textes littéraires ou des questionnaires et discours en sociologie et en psychologie sont intéressantes pour le traitement des textes scientifiques. On peut rendre hommage aux statisticiens Lebart et Salem d'avoir rassemblé les divers représentants de ces disciplines pour parler de "statistique textuelle" (noter le singulier) [Lebart et Salem, 1994]. Cette approche restaure une universalité des méthodes d'analyse des textes, seule manière selon nous de les valider.

## II-4-2 DE LA BIBLIOMETRIE DISTRIBUTIONNELLE A LA BIBLIOMETRIE DES REFERENCES BIBLIOGRAPHIQUES

### II-4-2. 1 Les lois bibliométriques :

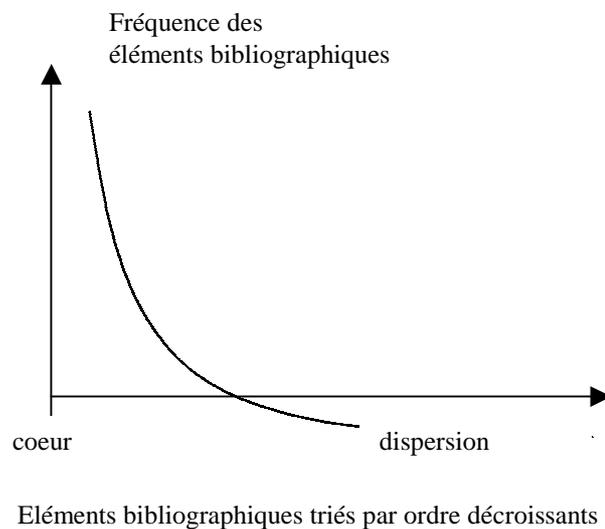
On parle de distribution statistique chaque fois que des données sont réparties sur un axe discret ou continu. Ces données sont issues de l'observation de fréquences d'événements, appelées généralement distributions bibliométriques. Nous pouvons citer celles relatives à :

- la productivité des revues scientifiques, mesurée en articles sur un sujet donné (loi de Bradford)
- la répartition des mots dans un texte (loi de Zipf)
- la production des chercheurs (loi de Lotka) : variation du pourcentage d'auteurs en fonction du nombre d'articles publiés
- l'utilisation des revues d'une bibliothèque

Ces distributions sont l'objet de ce que l'on appelle "la bibliométrie distributionnelle", prolongement de l'ancienne "bibliographie statistique".

a) Les trois premières lois reposent sur un type de distribution de forme cœur/dispersion. Le cœur représente le groupe des éléments les plus fréquents au sein des documents sélectionnés. La dispersion représente les nombreux éléments de basse fréquence au sein des mêmes documents.

Fig. 1 Cœur et dispersion d'une distribution bibliométrique



### **La loi de BRADFORD :**

Bradford considérait que l'activité de gestionnaire de bibliothèque est soumise au "chaos documentaire" de la littérature. L'un des problèmes qu'il pose peut s'énoncer ainsi : il n'est pas possible de s'abonner à tous les périodiques concernant un domaine, donc il faudrait trouver un moyen de sélectionner ceux qui sont les plus représentatifs d'un domaine. Il constate que peu de revues publient l'essentiel des articles sur un sujet. Il a cherché à ranger les périodiques en "zones" dégressives de productivité en fonction de leur proportion d'articles sur un sujet donné.

La loi de Bradford est une courbe représentant la somme cumulée des revues par rapport au logarithme de la somme cumulée des revues. Elle aboutit à la détermination de zones dont la première définit le noyau des meilleurs producteurs concernant le sujet.

### **Loi de LOTKA :**

Lotka cherche à déterminer la contribution de chaque chercheur à la science. Son domaine est la chimie.

Il a proposé une distribution du pourcentage d'auteurs par rapport au nombre d'articles publiés.

Très peu d'auteurs ont publié un grand nombre d'articles (le cœur) et une grande variété d'auteurs ont très peu publié.

### Loi de ZIPF :

L'auteur s'interroge sur la fréquence à laquelle les mots apparaissent dans un texte (ZIPFT 1949). Il comptabilise les occurrences des termes de l'Ulysse de Joyce, les classe par ordre décroissant de fréquence puis affecte à chaque mot un rang en commençant par 1 pour le mot le plus fréquemment apparu.

Il en résulte ce tableau :

*Tableau 18 : Loi de Zipf*

Rang	Fréquence	Produit rang*fréquence
10	2653	26530
30	926	27780
50	556	27800
100	265	26500
300	84	25200
500	50	25000
1000	26	26000
3000	8	24000
5000	5	25000
10000	2	20000
20000	1	20000
29899	1	29899

Il en déduit que le produit du rang par la fréquence est proche d'une constante. Pour notre part nous n'avons jamais vérifié expérimentalement cette égalité.

Certains auteurs ont cherché à exploiter cette distribution en tentant de découper des zones plus fines que le cœur et la dispersion qui mettraient en évidence une zone de bruit, une zone d'information intéressante, une zone d'information triviale [Quoniam, 1992].

b) une autre série de travaux porte sur les distributions relatives à la circulation de l'information [Morse, 1968 ; Ducasse, 1978 ; Lafouge, 1991]. Lafouge, à la suite de Ducasse, définit ainsi la bibliométrie :

*"Par donnée bibliométrique il faut entendre toute information de nature quantitative normalisée ayant trait aux différents éléments du système documentaire (documents, usagers,*

*transactions...)* qui après avoir été saisie, donne aux responsables un état actualisé dont l'analyse engendrera la décision."

L'étude de ces distributions relatives à la circulation est liée au facteur temps.

Ainsi Morse [1968] par exemple utilise des résultats de la recherche opérationnelle sur les processus markoviens pour mettre en place dans une bibliothèque un système d'analyse prévisionnelle de la demande d'ouvrages.

## **II-4-2. 2 - La création de l'ISI : de la gestion de la documentation à l'évaluation de la recherche**

Les années 1960 marquent à la fois la création de l'Institute for Scientific Information par Gardfield à Philadelphie et la naissance de la "scientométrie" (qui se distingue ainsi, pour une part, de l'ancienne bibliométrie).

Une nouvelle technique d'évaluation de l'activité scientifique, fondée sur l'étude des citations que se distribuent les auteurs d'articles, naît dans cet institut (l'ISI). Gross P.L.K. et Gross E.M., en 1927, entreprirent de comptabiliser non plus les documents mais les citations que les chercheurs faisaient, dans leur propre document, des travaux précédemment publiés. Ils effectuèrent ces comptages pour des journaux cités qui touchaient aux domaines de la chimie, puis rangèrent ces journaux par ordre décroissant du nombre de citations reçues. Ils venaient d'établir la liste des journaux qu'ils considéraient comme indispensables en chimie. Nous pouvons remarquer dans cette étude que le dénombrement est toujours basé sur le comptage d'articles, ceux-ci étant regroupés par catégories : l'appartenance aux mêmes journaux.

Gardfield a eu l'idée de constituer un répertoire ayant une couverture interdisciplinaire et qui regrouperait uniquement les articles publiés par les principaux périodiques scientifiques. ce "cœur" des revues étant déterminé par le taux de citations dont elles font l'objet.

L'activité principale de l'ISI est donc de collecter les articles publiés dans les revues les plus prestigieuses du monde pour toutes les branches de la science. Gardfield avait déjà imaginé en 1955 le principe du SCI. La première édition papier du SCI est parue en 1963 et couvrait la littérature mondiale de 1961. Elle s'étendait à 613 revues et contenait 1.4 millions de citations en 5 volumes. Depuis 2 répertoires concernant les autres branches de la science sont venus se joindre au SCI : le Science Social Citation Index publié depuis 1973 soit 1400 autres périodiques, et depuis 1978, Arts et Humanities Citation Index (A&HCI).

- **SCI :**

- le Citation Index (répertoire des citations par noms d'auteurs)
- le Source index (répertoire des publications par noms d'auteurs)
- le Permutem Subject Index (répertoire des mots du titre des publications)

- le **JCR** (depuis 1976)

- nombre de citations reçues par une revue (c'est-à-dire les citations des articles publiés par une revue).
- le facteur d'impact (FI) de chaque revue
- le classement des revues par thèmes (128) selon leur valeur de FI
- les 15 revues les plus citées pour chaque revue
- les 15 revues qui citent le plus chaque revue

- Base de données sur la distribution du nombre de publications par thème et par pays récemment mise à disposition par l'ISI [Okubo, 1996] pour une exploitation de cette base.

A tous ces répertoires fortement critiqués et critiquables s'ajoute, grâce à l'exploitation à grande échelle des références bibliographiques rendue possible par le développement des technologies de l'information et de la communication, tout un travail, tout aussi critiquable quant à sa valeur, autour des indicateurs destinés à évaluer l'activité scientifique et technique.

### II-4-2. 3 Les indicateurs

Ils sont basés sur 2 types de dénombrements :

- Le comptage des références où est présent l'élément bibliographique
- Le comptage des citations reçues : nombre d'articles qui citent l'élément bibliographique.

Voici quelques exemples d'indicateurs bibliométriques :

a) L'évaluation des revues

- Le facteur d'impact de l'ISI

Nombre de citations reçues l'année x pour les articles publiés par une revue pendant les deux années précédentes / nombre d'articles publiés par cette même revue pour ces deux années précédentes.

$$C_x / P_{x-1} + P_{x-2}$$

- L'indice d'immédiateté

nombre de citations reçues l'année y pour les articles la même année par une revue (par exemple nombre de fois que l'on cite en 1997 les articles de *Jasis* de 1997) / le nombre d'articles publiés cette année-là.

b) L'évaluation des domaines d'activités des pays

- L'indice d'avantage

M nombre total de publications tous pays et tous domaines confondus

$M_{pd}$  nombre de publications réelles pour le pays p dans ce domaine d

$M_p$  nombre de publications du pays p concerné

$M_d$  nombre de publications du domaine étudié

$$\frac{M_{pd}}{M_p \cdot M_d / M}$$

- L'indice d'activité [Barré 1991, Callon et Leydesdorff 1987]

$$\frac{M_{pd} / M_d}{M_p / M}$$

Si la valeur de l'indice est inférieure à 1 on en déduit que le pays a une contribution faible dans le domaine, et inversement si l'indice est supérieur à 1.

De sérieuses difficultés sont liées aux comptage des citations :

- les citations erronées qui renvoient à des sources secondaires plutôt qu'à l'auteur principal de la découverte.
- l'auto-citation

- la différence de nature entre les citations : certaines sont faites dans le contexte de la critique. Mais il reste qu'elles manifestent un fait de lecture de l'article ou de l'auteur cités.

- on cite davantage les revues prestigieuses.

- la variation de la pratique de la citation dans chaque discipline

- le taux de citation dépend du type de document (article de fond ou synthèse).

Par exemple les articles de méthodologie seraient plus cités car les auteurs, pour éviter de décrire la méthode qu'ils utilisent, citent les documents où elle est expliquée.

- un auteur cite plus facilement ses compatriotes

- le taux de citation varie selon la nationalité des auteurs

- les auteurs d'articles cherchent à faire référence à des articles de chercheurs renommés afin de mieux convaincre de la solidité de leur argumentation

- les auteurs-clés à l'origine du domaine peuvent ne pas apparaître.

Le nombre de citations pris comme une donnée quantitative pour "fabriquer" un indicateur peut être soumis à de nombreuses critiques liées aux remarques faites ci-dessus.

**Là est la limite de tous ces indicateurs** ; d'où l'intérêt des méthodes relationnelles (qui nécessitent l'utilisation de méthodes statistiques complexes : analyse de données, statistique multidimensionnelle) pour repérer des structures qualitatives entre les données.

## II-4-2. 4 L'analyse des associations

### II-4-2. 4,1 : Les objets

Les objets que l'on cherche à associer sont :

- les cocitations (d'articles, d'auteurs, de revues) afin de repérer, par exemple, les "collèges invisibles" d'auteurs c'est-à-dire les auteurs travaillant ensemble ou se référant les uns aux autres.

- les mots associés (nous la décrivons ci-après)

- la co-occurrence de codes documentaires : il s'agit de mettre en évidence des associations de codes, par exemple des codes de la classification internationale des brevets.

- relation pays / domaines (en utilisant les données de l'ISI par exemple) afin de comparer le profil scientifique de différents pays ou bien l'évolution du profil scientifique d'un pays<sup>34</sup>

#### II-4-2. 4, 2 Les méthodes

Elles relèvent de l'analyse de données. L'analyse factorielle et les classifications automatiques sont les deux grandes familles de méthodes utilisées pour la mise en évidence des relations entre les objets que nous avons listés ci-dessus.

Nous empruntons à Lebart et Salem [1994] l'énoncé du principe commun de ces méthodes : *"Chacune des dimensions d'un tableau rectangulaire de données numériques permet de définir des distances (ou des proximités) entre les éléments de l'autre dimension. Ainsi l'ensemble des colonnes (qui peuvent être des variables, des attributs) permet de définir à l'aide de formules appropriées des distances entre lignes (qui peuvent être des individus, des observations). De la même façon, l'ensemble des lignes permet de calculer des distances entre colonnes.*

*On obtient ainsi des tableaux de distances, auxquelles sont associées des représentations géométriques complexes décrivant des similitudes existant entre les lignes et entre les colonnes des tableaux rectangulaires à analyser."* [Lebart et Salem 1994.p.80]

Tandis que les méthodes factorielles produisent des représentations graphiques où les associations statistiques sont traduites par des proximités géométriques entre points-lignes et points-colonnes, les méthodes de classification opèrent des regroupements en classes ou en famille de classes hiérarchisées des lignes et des colonnes. Ces deux types de méthodes peuvent non seulement être utilisées sur les mêmes tableaux de données mais aussi se compléter.

---

<sup>34</sup> Les études sur ce sujet sont nombreuses. On peut citer, à titre illustratif : Okubo [1996] ; Doré *et alii* [1996].

### a) les classifications

Elles peuvent être :

- non hiérarchiques : c'est le cas de la classification à lien simple de la méthode des mots associés.

- hiérarchiques telles la Classification Ascendante Hiérarchique (CAH) ou la Classification descendante Hiérarchique (CDH). La CAH peut être appliquée sur des tableaux de présence/absence ou sur les coordonnées de l'analyse factorielle pour une autre présentation des résultats c'est-à-dire des liens.

Les opérations sont les suivantes :

- choix d'un indice de similarité ou d'une distance. Le but est de mesurer l'intensité d'un lien en fonction de la présence simultanée ou de l'absence simultanée des éléments à mettre en relation etc.

- pairage. Tableau de fréquence des paires

- Tri. classement par ordre décroissant

- constitution d'agrégats.

Voici un exemple d'utilisation d'une méthode de classification non hiérarchique dans le cadre de la scientométrie. Il s'agit d'une méthode de mots associés appliquée par le logiciel LEXIMAPPE issu du Centre de sociologie de l'Innovation de l'Ecole des Mines de Paris.

Les principales opérations effectuées et leurs résultats sont :

- Constitution d'une matrice mots \* documents

	M1	M2	Mi	Mj
M1500				
D1			1	1
D2			1	
Di			1	
Dj			1	
Dm				

- Transformation en matrice carrée mots \* mots
- Calcul de l'association des paires avec l'indice :

$$E_{ij} = C_{ij}^2 / C_i \cdot C_j$$

que l'on peut écrire :

$$E_{ij} = (C_{ij} / C_i) * (C_{ij} / C_j)$$

Sachant que :

- $C_{ij}$  représente le nombre de fois que le mot  $i$  et le mot  $j$  co-occurrent (c'est-à-dire apparaissent simultanément) dans le même document.
- $C_i$  et  $C_j$  sont respectivement le nombre de fois que chaque mot apparaît.

En d'autres termes c'est la probabilité d'avoir le mot  $i$  quand on a le mot  $j$  multipliée par la probabilité d'avoir le mot  $j$  quand on a le mot  $i$  :

$C_i \cdot C_j$  correspond donc à l'effectif théorique (cas de l'indépendance des mots) et  $C_{ij}$  à l'effectif observé.

Comme l'explique Courtial [1990], l'indice Leximappe " est une sorte de mesure de la relation "et" entre les deux mots  $i$  et  $j$  au sens où l'un doit impliquer l'autre puisque le coefficient est faible si l'un des deux termes de la multiplication est faible."

Ce type de méthode ne fait pas intervenir la mesure de toutes les associations possibles entre mots-clés contrairement à l'analyse factorielle des correspondances. Elle ne raisonne pas en terme d'espace mais en termes de **chaînes de mots-clés**.

Le calcul des paires et de l'intensité des liens conduit à la construction d'agrégats.

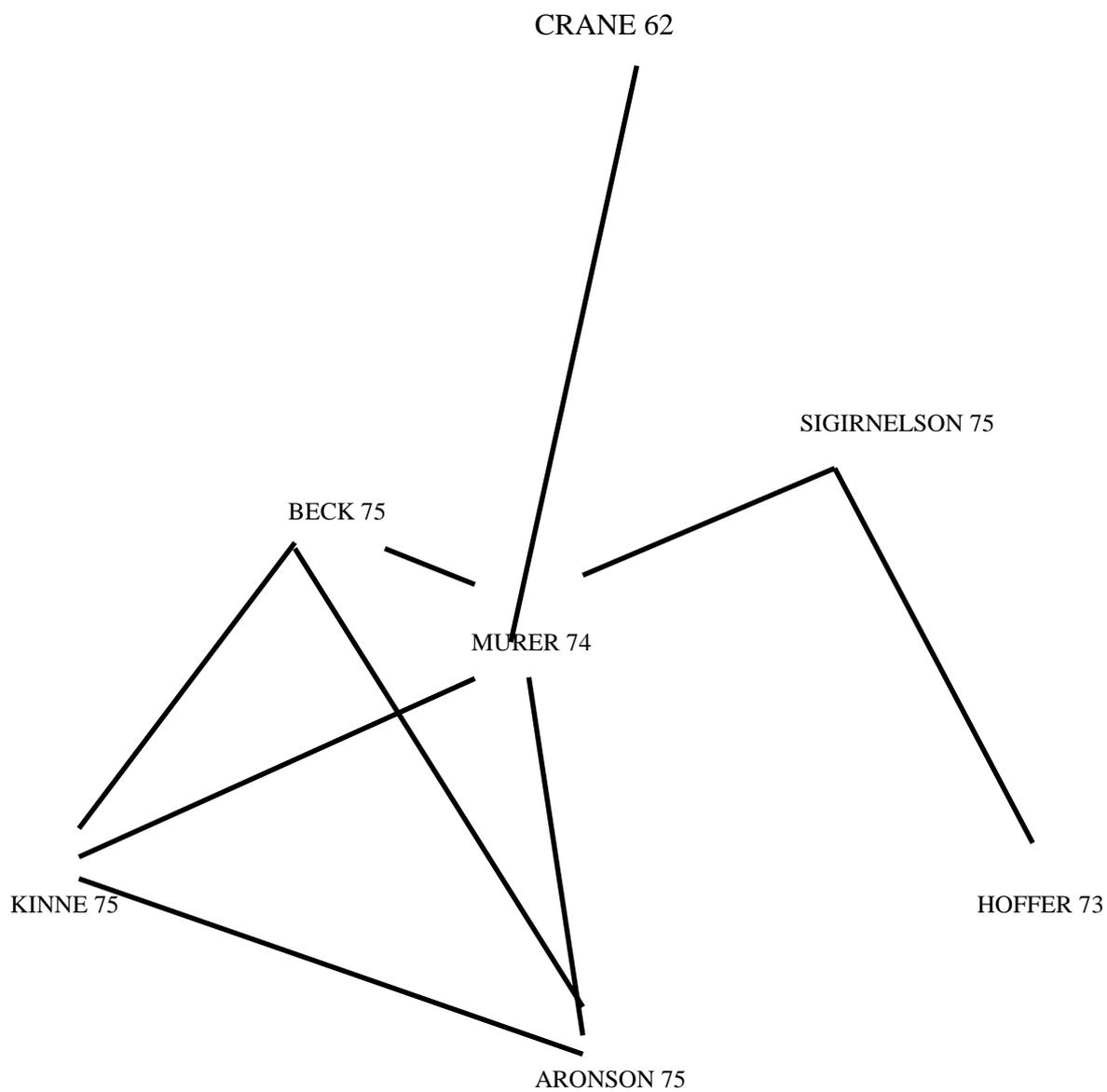
- Après avoir calculé le coefficient d'association entre mots-clés, les paires de mots-clés sont rangés par ordre décroissant. En parcourant cette liste depuis le début le programme construit les doublets, puis les triplets, etc. de mots associés de façon à fournir un graphe n'excédant

pas dix mots. A chaque fois que l'on obtient un graphe de 10 mots-clés, il est imprimé et les mots sont retirés de la paire restante de la liste.

- On obtient ainsi une liste de clusters telle que le cluster ait des liens maximaux. Un cluster peut se définir par sa valeur moyenne des liens internes qu'on appelle **densité**. Il peut se définir aussi par la valeur moyenne de ses liens externes, c'est-à-dire des liens entre mots qui le constituent et les mots d'autres clusters qu'on appellera **centralité**.
- Disposition des clusters sur un plan selon un système à deux axes perpendiculaires : c'est le **diagramme stratégique**. Un axe répartit les agrégats selon leur densité (mesure de la cohésion interne de l'agrégat) et l'autre selon la centralité c'est-à-dire les liens qu'ils entretiennent avec les autres agrégats. Chaque agrégat est symbolisé sur les graphes par un des mots-clés du groupe sélectionné par un indice.
- L'idée de cycle de vie d'un thème est abordée par l'équipe du CSI de l'Ecole des Mines dans "Gestion stratégique de la recherche et de la technologie..." [Callon, Larédo, Mustar, 1995]. On voit apparaître un indice de transformation [*ibid*, p.209] qui mesure les modifications du thème au cours du temps : plus l'indice est élevé plus le contenu se modifie. Outre cet indice, une autre méthode pour appréhender les transformations d'un thème au cours du temps consiste à comparer les différents diagrammes stratégiques obtenus dans un temps  $t$ ,  $t + 1$  etc. pour voir comme un thème passe d'un endroit à un autre du diagramme. C'est cette méthode qui est utilisée dans l'exemple que nous donnons ci-après.

Figure 2

GRAPHE DE CO-CITATIONS



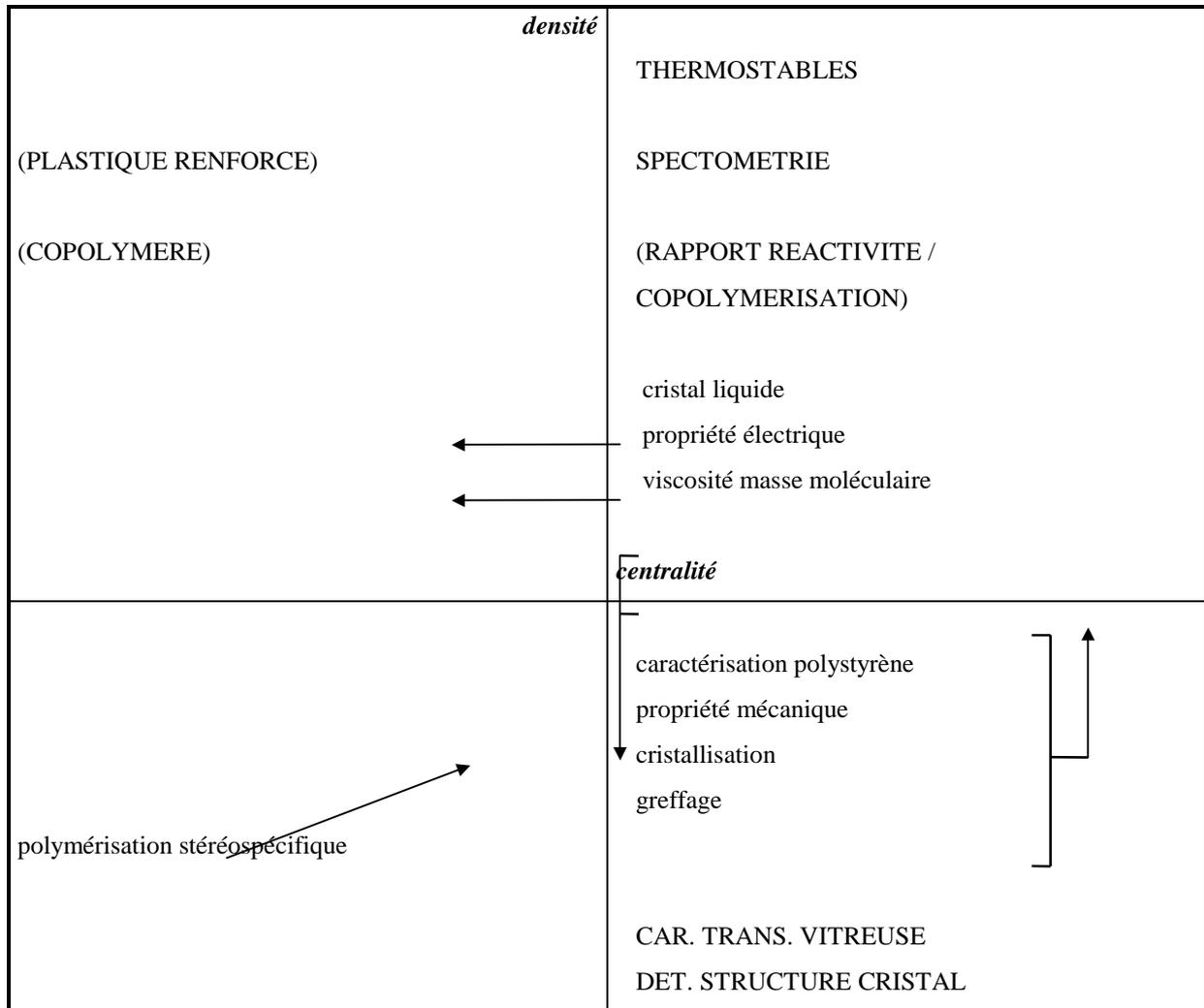
**TRANSPORT DANS LES VESICULES MEMBRANAIRES DE LA BORDURE EN BROSSE DU REIN**

D'après Courtial [1990, p. 54)

Figure 3

**DIAGRAMME STRATEGIQUE :**

carte simplifiée des déplacements des thèmes de 1973 à 1986



Les thèmes en capitale ne se déplacent pas. S'ils ne sont pas entre parenthèses, ils sont toujours présents.

D'après Courtial [1990, p. 106]

## b) L'analyse factorielle

Dans l'espace à  $n$  dimensions sont repérées les directions privilégiées de l'étalement du nuage dans l'espace qui constituent les facteurs de l'analyse. C'est l'organisation des points dans la totalité de leurs associations qui est soumise à l'interprétation facteur par facteur.

### II-4-3 DE LA SCIENTOMETRIE A L'ANALYSE DE CONTENU

La question qui se pose maintenant est celle du traitement du texte libre. L'objectif est l'analyse du contenu de bases documentaires.

L'analyse de contenu au sens où nous l'entendons se distingue de celle de ses fondateurs de l'école de Chicago. Elle garde l'objectif qui consiste à s'intéresser au "*contenu manifeste de la communication*" [Bereleson et Lazarfeld, 1948] mais élimine la méthode qui suppose la définition de classes *a priori* à l'intérieur desquelles se font les décomptes d'éléments textuels.

Les méthodes statistiques sont les mêmes pour la bibliométrie des références bibliographiques et pour l'analyse de contenu. Cependant le niveau de complexité n'est pas le même et la question est de savoir dans quelle mesure leur application au texte libre est rigoureuse.

Le principe général est le suivant :

On utilise une distance ( $d$ ) ou un indice de similarité ( $I_s$ ) sachant que :

$$1 - I_s = d$$

Ce calcul repose toujours sur la comparaison entre effectif observé ( $O$ ) et effectif théorique ( $T$ ) d'une des données listées ci-dessous : la distance du  $\chi^2$   $[(O-T)^2/T]$  comme l'indice Leximappe  $[O^2/T]$  reposent sur ce type de rapport <sup>35</sup>.

Les données numériques utilisées sont :

- le rapport entre la fréquence du mot dans un document et la fréquence du mot dans la base.
- la fréquence du mot dans l'unité de contexte.
- la fréquence d'association

Les méthodes utilisés par la scientométrie (l'analyse des associations) ont été récemment appliquées au texte libre (voir par exemple le logiciel SAMPLER). Mais elles présentent de sérieux inconvénients : le nombre de paires explose et les clusters constitués sont arbitrairement coupés.

### **II-4-3. 1 Les classifications hiérarchiques**

Nous avons fait peu d'usage des classifications hiérarchiques de mots dans nos expérimentations en documentation bien que nous les ayons systématiquement calculées. Nous notons toutefois l'existence de travaux utilisant un type de classification hiérarchique sur le texte intégral lemmatisé<sup>36</sup>. Les travaux de Max Reinhert [1987 ; 1993 entre autres] se situent dans cette perspective. L'auteur a mis au point une méthode de classification descendante hiérarchique des unités de contextes des textes divers (œuvre littéraire, entretiens, récits de cauchemar etc.).

### **II- 4- 3. 2 L'analyse factorielle**

Parmi les différentes méthodes d'analyse factorielle, l'une a des propriétés intéressantes pour le traitement des données linguistiques. Il s'agit de l'analyse factorielle des correspondances. C'est cette méthode que nous utiliserons dans nos expérimentations (voir Partie III de cette étude).

#### II- 4- 3. 1, 1 Analyse Factorielle des Correspondances :

*L'analyse des correspondances est une technique de description de tableaux croisés (ou tables de contingence) ou de tableaux binaires de type "présence-absence". Deux ensembles (individus et variables, ou observations et variables, ou mots et textes) sont mis en correspondance sous la forme de tableaux rectangulaires de données numériques avec, à l'intersection de la ligne et la colonne, le nombre de fois que l'élément est présent, ou bien l'indication par 0 et 1 de sa présence ou de son absence.*

Nous présentons la méthode en insistant uniquement sur les trois points qui ont une importance particulière pour le traitement des données linguistiques : la notion de profil, le calcul des proximités (la distance utilisée) et la représentation simultanée. Nous reprenons les formules de Lebart et Salem 1994.

---

<sup>35</sup> Voir p.86.

<sup>36</sup> Lemmatiser revient à ramener à une forme unique les différentes formes d'un terme (singulier/pluriel, féminin/masculin, etc.)

### a) Présentation de ces trois points :

#### 1- la notion de profil

Le voisinage entre deux formes graphiques signifie le voisinage des profils (lignes ou colonnes).

On notera  $f_{ij}$  le terme général de la table des fréquences dont les éléments sont préalablement divisés par l'effectif total  $k$ .

Cette table possède  $n$  lignes et  $p$  colonnes

$f_i$  désigne la somme des éléments de la ligne  $i$

$f_j$  désigne la somme des éléments de la colonne  $j$  de cette table.

Le profil de la ligne  $i$  est l'ensemble des  $p$  valeurs :

$$\frac{f_{ij}}{f_i}, j = 1, \dots, p$$

Le profil de la colonne  $j$  est l'ensemble des  $n$  valeurs :

$$\frac{f_{ij}}{f_j}, i = 1, \dots, n$$

#### 2- Les proximités (distance du $X^2$ )

La distance entre deux points-lignes  $i$  et  $i'$  est donnée par la formule :

$$d^2(i, i') = \sum_{j=1}^p (1/f_j) (f_{ij}/f_i - f_{i'j}/f_{i'})^2$$

La distance entre deux points-lignes  $j$  et  $j'$  est donnée par la formule

$$d^2(j, j') = \sum_{i=1}^n (1/f_i) (f_{ij}/f_j - f_{ij'}/f_{j'})^2$$

### 3- La représentation simultanée :

Si  $\phi_i$  désigne la coordonnée d'un point-ligne  $i$  sur le premier axe factoriel et si  $\Psi_j$  désigne celle d'un point-colonne  $j$  sur ce même axe, on a le système de relations remarquablement symétriques :

$$\phi_i = \beta \sum_{j=1}^p (f_{ij} / f_{i.}) \psi_j$$

$$\psi_j = \beta \sum_{i=1}^n (f_{ij} / f_{.j}) \phi_i$$

Le coefficient  $\beta$  est un coefficient positif supérieur à 1 rendant possibles les relations 1 et 2 simultanément ; il compense l'effet contractant d'une représentation qui serait doublement barycentrique.

D'autre part, les coordonnées sont centrées c'est-à-dire vérifient les relations :

$$\sum_{i=1}^n f_{i.} \phi_i = 0$$

#### **b) L'hypothèse linguistique de l'AFC**

$$\sum_{j=1}^p f_{.j} \psi_j = 0$$

Bien qu'elle s'applique à des données de multiples disciplines, l'analyse des correspondances a ceci de spécifique qu'elle repose sur une notion de linguistique : l'équivalence distributionnelle. Cette méthode évacue l'analyse syntaxique ou plutôt la nécessité d'un découpage grammatical comme moyen de l'analyse sémantique d'un texte.

Benzécri ramène à l'étude des langues mortes, l'origine des théories distributionnalistes. Le linguiste recense les signes de fréquence de longueur inégale, dont les textes lui paraissent composés ; remarque des groupes de signes souvent employés qui peuvent être des mots ou des locutions usuelles ; note que certaines unités ne sont ordinairement suivies que de certaines autres appartenant à une liste assez courte... L'étude muette des cooccurrences, des rencontres d'éléments, qui est tout ce que permet au linguiste un corpus isolé, peut aussi être

faite sur une langue que l'on comprend, mais dont on s'abstient de considérer le sens. C'est cette idée qui conduit Benzécri, à défendre une méthode inductive en linguistique. Dans cet objectif il exploite la définition de l'analyse distributionnelle de Harris (dans : *Distributionnal Structure in Word*, 1954) : "Par distribution d'un élément, on entend l'ensemble de ses contextes possibles". En toute rigueur, le contexte d'un mot c'est le texte moins ce mot. Or à l'échelle de la langue, l'ensemble des textes ou des phrases dans lesquels peut s'insérer un mot est infini et à l'échelle d'un corpus donné, si ample soit-il, on ne trouvera de cette diversité de contextes possibles que quelques cas qui n'en représentent pas un échantillon statistique suffisant (c'est le problème du modèle probabiliste selon Benzécri comme nous le verrons). C'est pourquoi les dénombrements s'opèrent sur des contextes schématiques supposant un choix et c'est là l'une des difficultés de l'AFC : la sélection des unités de contexte dans lesquelles se feront les comptages et le calcul des profils.

Qu'est-ce que l'équivalence distributionnelle de données linguistiques ?

Reprenons l'exemple de Benzécri .

Soit un corpus  $k$  dont toutes les phrases se composent d'un nom (sujet) suivi d'un verbe (prédicat) : dans chacune de ces phrases le contexte du nom n'est que le verbe dont il est sujet ; et le contexte du verbe est le nom qui est son sujet. La distribution d'un nom sera donc l'ensemble des verbes dont il se trouve avoir été sujet ; et de même pour la distribution d'un verbe. De façon précise, notons  $I$  l'ensemble des noms et  $J$  l'ensemble des verbes rencontrés dans le corpus  $k$ . Toute l'information statistique apportée par  $k$  peut s'inscrire dans un tableau rectangulaire dont  $I$  est l'ensemble des lignes et  $J$  celui des colonnes : à l'intersection de la ligne  $i$  et de la colonne  $j$  on place l'entier  $k(i, j)$ , nombre de fois que, dans le corpus  $k$ , figure la phrase " $ij$ ". Ainsi la distribution du nom  $i$  n'est autre que la ligne  $i$  du tableau des  $k(i, j)$ , c'est-à-dire la suite  $\{k(i, j) \mid j \in J\}$ , des nombres de fois que  $i$  est sujet de chacun des verbes ; et de même la colonne  $j$  donne la distribution du verbe  $j$ .

Tableau 19 : l'équivalence distributionnelle

	aboyer	dormir	manger	miauler	ronfler	voler
avion	0	0	0	0	3	7
chacal	1	1	1	0	1	0
chat	0	5	3	4	1	0
chien	3	3	3	0	3	0
coyote	2	1	1	0	1	0
moteur	1	0	3	0	8	0
oiseau	0	2	3	0	0	6

On pourrait aussi, à partir de ces données, étudier un tableau de compatibilité logique (de présence/absence) comme celui-ci :

**Tableau 20 : Equivalence distributionnelle suite**  
**tableau de compatibilité logique**

	aboyer	dormir	manger	miauler	ronfler	voler
avion	0	0	0	0	1	1
chacal	1	1	1	0	1	0
chat	0	1	1	1	1	0
chien	1	1	1	0	1	0
coyote	1	1	1	0	1	0
moteur	1	0	1	0	1	0
oiseau	0	1	1	0	0	1

Sur le premier tableau on observe la proximité distributionnelle de l'avion et plus encore de moteur aux êtres vivants : un moteur en effet ronfle, mange (du combustible) et parfois aboie (quand il a des ratés). Sur le tableau de compatibilité on voit que chien et chat ne se distinguent que par leur voix. Sur ce même tableau chien, chacal, coyote coïncident. De plus, quant aux fréquences, les lignes chien et chacal apparaissent exactement proportionnelles : bien qu'en valeur absolue "*chien*" soit trois fois plus employé que chacal (12 contre 4), en valeur relative leurs emplois sont les mêmes :

chacal	1/4	1/4	1/4	0	1/4
chien	1/4	1/4	1/4	0	1/4

On dira que chacal et chien sont des synonymes distributionnels, ou encore qu'ils sont distributionnellement équivalents. C'est cette équivalence distributionnelle qu'exprimeront les relations de proximité d'une l'AFC faite sur un tel tableau mais en y ajoutant l'ensemble des relations de l'individu à son groupe et à celui auquel il est comparé ; c'est-à-dire en permettant une représentation multi-dimensionnelle. C'est pourquoi l'AFC utilise la distance du  $\chi^2$  (chi2) appelée encore distance distributionnelle qui respecte la propriété suivante : si deux lignes du tableau brut sont proportionnelles et qu'on les remplace par une seule ligne qui en soit la somme, alors la distance entre colonnes ne change pas ; le même résultat est obtenu avec les lignes pour des colonnes proportionnelles.

Certaines méthodes consistant à quantifier des proximités entre termes à partir de leur identité distributionnelle dans des relations syntaxiques précises <sup>37</sup> (relations sujet/prédicat ; sujet-verbe-objet). Elles exploitent l'équivalence distributionnelle dans des contextes restreints qui nécessitent leur reconnaissance et leur extraction puis l'utilisation de méthodes de classification fondées sur des mesures de similarité. La différence majeure entre les deux approches est que l'une (celle fondée sur les similarités) considère le contexte comme un *ordre* de dépendance syntaxique alors que la seconde casse tout lien de dépendance entre syntaxe et sémantique

L'hypothèse distributionnaliste en linguistique que reprend à son compte Benzécri pour justifier l'utilisation de l'AFC en linguistique s'oppose avant tout au modèle probabiliste. Ce modèle peut être résumé de la façon suivante : si au cours d'une suite de tirages aléatoires de documents, la fréquence d'apparition du mot dans chacun des documents reste stable (c'est-à-dire qu'il n'y a pas de différence significative entre les fréquences), alors le mot n'est caractéristique d'aucun document ; si pour un des tirages, la fréquence est anormalement élevée (significativement supérieure) alors ce mot est une des caractéristiques de ce document.

---

<sup>37</sup> Voir L.Favier [1994].

C'est précisément à l'idée que l'insertion de chaque mot fait l'objet d'un acte indépendant et constitue à ce titre une unité statistique que s'oppose le distributionnalisme de l'A.F.C.

Pour être plus complet sur ces aspects de théorie linguistique relatifs à l'AFC il faut encore mentionner le contexte historique dans lequel cette méthode a pris son essor. A cet effet, nous citons les mots de Benzécri dans l'avant-propos de l'ouvrage de 1980 :

*"Vers 1960, la traduction automatique semblait un objectif assez rapidement accessible (sont apparus depuis des obstacles qu'on n'a pas surmontés...). L'Association pour la Traduction Automatique A.T.A.L.A, fondée à l'initiative de E. Delavenay aidait efficacement aux rencontres des chercheurs français ; (...). La linguistique mathématique (...) était alors dominée par le renom de N. Chomsky dont le petit volume Syntactic Structures s'imposait à tous.*

*Entre autres thèses, N. Chomsky affirme là, qu'il ne peut exister de procédure systématique pour déterminer la grammaire d'une langue ou plus généralement les structures linguistiques, à partir d'un ensemble de données tel qu'un recueil de textes que les linguistes nomment corpus (...). Contre cette thèse (idéaliste, en ce qu'elle tendait à séparer le jeu de l'esprit des faits qui en sont l'inspiration et l'objet) on a voulu proposer, à défaut d'un algorithme universel pour passer de 10.000 pages de texte d'une langue à une syntaxe doublée d'une sémantique, une méthode inductive efficace pour traiter utilement des tableaux de données qu'on pouvait immédiatement recueillir ; avec à l'horizon l'ambitieux étagement des recherches successives ne laissant rien dans l'ombre des formes, du sens et du style."*

On voit resurgir ici implicitement, à travers l'opposition à Chomsky, l'idée selon laquelle le sens d'un texte est indépendant des structures syntaxiques qu'il met en œuvre. D'où l'idée que le traitement automatique de la langue peut se passer de l'analyse syntaxique en utilisant une méthode strictement inductive, ne partant d'aucun modèle ni d'aucune hypothèse mais seulement des tableaux de données, à l'image du décryptage des langues anciennes et mal connues. Alors seule la distribution des éléments constituants peut donner une image de la cohérence du corpus traité.

L'AFC est toutefois une méthode peu utilisée en documentation. Une consultation de la base de données LISA ("*Library and Information Science Abstracts*") semble confirmer notre impression selon laquelle l'utilisation de l'AFC en science de l'information semble surtout limitée aux études scientométriques. Cela est d'autant plus surprenant que l'AFC avait bien été conçue au départ pour le traitement du contenu des textes et des travaux sur l'indexation automatique sont à noter au moins dès 1984 [HAMLAT 1984].

## II- 4- 3 1, 2 Latent semantic analysis (LSA) ; Latent Semantic Indexing (LSI)

### **a) introduction**

L'application de l'analyse factorielle en documentation a pris le nom aux Etats-Unis de "*Latent semantic analysis*" et "*Latent Semantic Indexing*". L'objectif de cette méthode est d'obtenir la structuration sémantique<sup>38</sup> d'une base de données en fonction des termes d'une requête, ce qui n'est pour nous que l'une des applications possibles de ce type de méthode en documentation.

La publication de travaux sur la LSA/LSI semble débiter à la fin des années 1980 avec, à notre connaissance, une première référence en 1988 dans les *Proceedings de l'ASIS* (American Society of Information Science). Ils émanent des chercheurs des laboratoires Bell (*Bell Communications Research*) et notamment de Dumais.

Leurs travaux ont donné lieu à deux dépôts de brevet :

- U.S. Patent N°. 4, 839, 853, June 1989 : "*Computer information retrieval using latent semantic structure*".
- U.S. Patent N°. 5, 301, 109, Apr. 5, 1994 : "*Computerized cross-langage document retrieval using latent semantic indexing*".

Il est à noter également que l'objectif de l'utilisation de cette méthode est de construire un moteur de recherche : le problème est posé en termes de sélection de documents en réponse à une requête.

Cela signifie que l'on ne cherche pas, dans cette perspective, à interpréter l'organisation dans son ensemble mais que l'on se concentre sur la délimitation d'un espace "de pertinence" contenant les documents à retourner à l'utilisateur. La stratégie que nous utilisons (voir IIIème Partie de notre étude), à savoir : constituer des bases de textes sur des sujets en utilisant des filtres grossiers (une chaîne de caractères, une expression, une équation logique) puis en

---

<sup>38</sup> L'article de Deewester *et alii* [1990] définit ainsi la notion de "structure sémantique": "*By semantic structure we mean here only the correlation structure in the way in which individual words appear in documents ; "semantic" implies only the fact that terms in a document may be taken as referents to the document itself or to its topic*".

dégager la structure thématique par une AFC pour en connaître le contenu est une démarche tout à fait étrangère, semble-t-il, au courant américain de la LSA.

Les exemples de textes traités par les équipes faisant de la LSA sont des titres, des résumés, des dossiers médicaux [Chute, 1995].

Les auteurs ont évalué leur méthode par rapport aux autres modèles de recherche d'information ainsi, Lochbaum [1989] pour la comparaison avec le modèle vectoriel, Swanson *et alii* [1990] pour la comparaison avec le modèle booléen, Chute [1995] pour une comparaison plus générale.

### **b) La méthode de LSA : Singular Value Decomposition : SVD**

La méthode proposée par les auteurs se fonde sur la décomposition en valeurs singulières d'une matrice rectangulaire due à Eckart et Young [1936] : "*The approximation of one matrix by another of lower rank*".

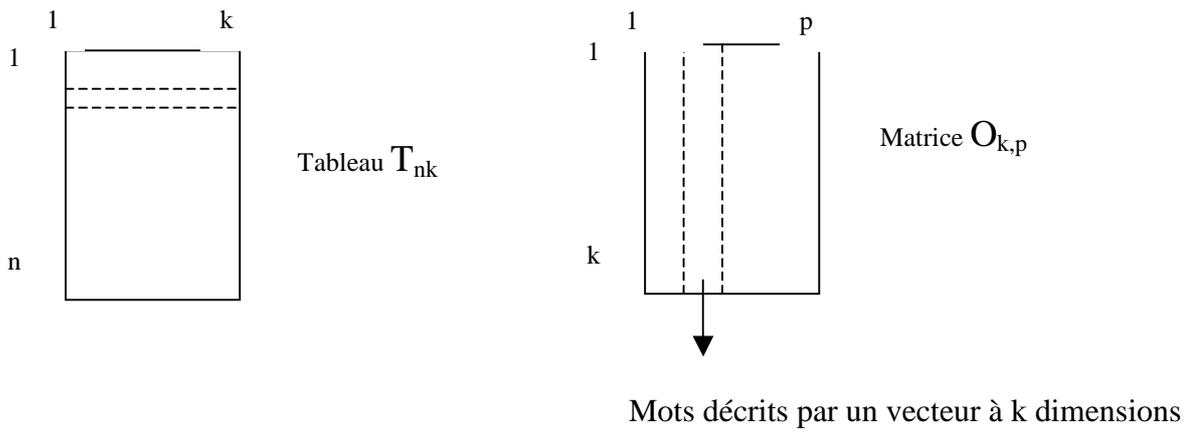
Un tableau rectangulaire  $X_{(n,p)}$  de terme générique  $X_{ij}$  peut se présenter sous la forme :

$$X_{(n,p)} = T_{n,r} S_{(r,r)} O_{r,p}$$

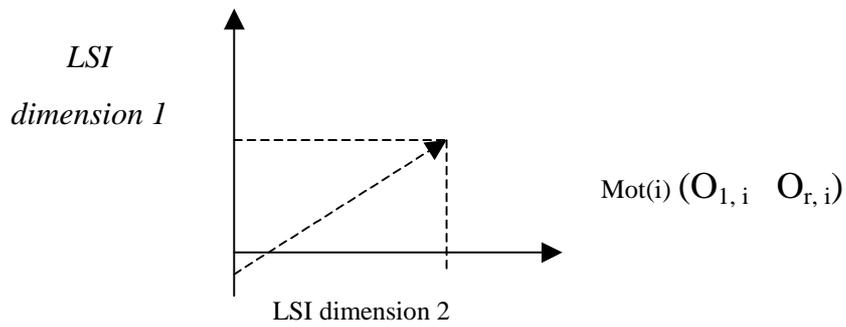
$S_{(r,r)}$  étant une matrice diagonale de rang  $r$  (matrice de valeurs propres) et une bonne approximation de la matrice  $X_{(n,p)}$  est obtenue en utilisant la sous-matrice diagonale  $S_{(k,k)}$  constituée des  $k$  plus fortes valeurs propres. On a alors :

$$X_{(n,p)} \approx T_{n,k} S_{(k,k)} O_{k,p}$$

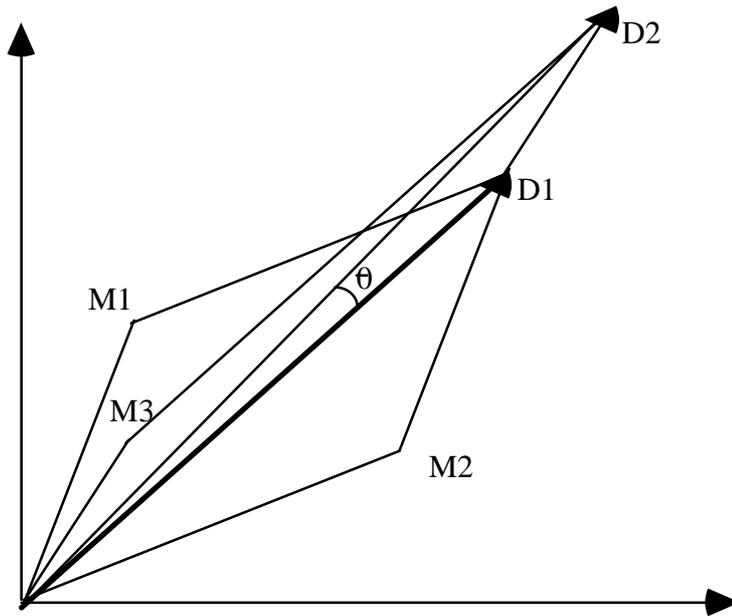
Si  $X_{(n,p)}$  est le tableau des correspondances documents\*mots, le tableau  $T_{n,k}$  représente la description des documents par des vecteurs à  $k$  dimensions tandis que le tableau  $O_{k,p}$  représente la description des mots par des vecteurs à  $k$  dimensions



La représentation géométrique des mots se fait alors comme suit (cas où  $k=2$ )



Un document sera représenté par la somme vectorielle pondérée des mots qui le constituent.



La proximité entre Document 1 et Document 2 est exprimée par le cosinus  $\theta$  de l'angle formé par les deux vecteurs D1 et D2.

En résumé la SVD est une méthode proche de l'analyse en composante principale non centrée non réduite. La différence avec l'AFC est donc le mode d'estimation des proximités : cosinus pour la SVD, proximité barycentrique en AFC.

Une requête est également représentée par un vecteur égal à la somme vectorielle des vecteurs mots qui constituent la requête.

### c) Présentation de Latent Semantic Analysis

Dans l'article de 1990 souvent cité comme l'article de référence en la matière [Deerwester *et alii* 1990], les auteurs présentent comme une nouvelle méthode d'indexation et de recherche d'information l'indexation par analyse sémantique latente. On cherche par cette méthode à mettre en évidence une structure sémantique c'est-à-dire une association de termes et de documents permettant de dégager un lien entre les termes de la requête et des documents de la base : ce lien serait synonyme de relation de pertinence. Les requêtes sont traitées comme des pseudo-documents et leurs termes se positionnent donc dans l'espace des mots et des documents de la base comme n'importe quel document. La position dans l'espace tient lieu de "nouvelle sorte d'indexation sémantique".

La technique utilisée de "*singular-value decomposition*" (décomposition en valeur singulière) décompose la matrice mots\*documents en facteurs orthogonaux, sur 100 facteurs dans Deerwester *et alii* [op.cit.], et les documents dont la valeur du cosinus par rapport à l'espace de projection atteint un certain seuil sont retournés à l'utilisateur.

Les méthodes d'indexation et de recherche d'information se heurtent à deux grandes difficultés que l'on peut résumer comme celles de **la synonymie et de la polysémie**, la synonymie entendue comme différentes façons de se référer à un même objet et la polysémie comme la possibilité pour un mot d'avoir plusieurs sens. Le problème de la synonymie (plusieurs manières d'exprimer la même chose) tend à diminuer le taux de rappel des documents : [Furnas, Launder, Gomez et Dumais, 1987] ont montré par exemple que dans moins de 20% des cas deux personnes choisissent le même mot pour désigner un objet bien connu. Quant à la polysémie (l'ambiguïté) elle tend à diminuer le taux de précision. Par exemple, des termes tels que : "chip" en anglais ou "état" en français ou "injection" ont de très nombreuses significations ("état" doit être distingué d' "Etat" sous peine de n'avoir plus le même sens, de même "injection" peut renvoyer à piqûre ou à technique de transformation des polymères).

Cela tient essentiellement à trois facteurs.

Le premier est que ni l'expansion automatique ou intellectuelle de la requête ni la construction d'un thesaurus ne sont des solutions satisfaisantes. Les termes ajoutés peuvent être polysémiques d'où une rapide dégradation de la précision [Spark-Jones, 1972]. L'interactivité n'améliore pas réellement les choses <sup>39</sup>

Le second facteur est la carence des méthodes automatiques pour traiter la polysémie.

Le troisième facteur est l'état des méthodes en matière d'indexation automatique et de système de recherche d'information. Dans ces systèmes chaque mot est traité séparément.

Les auteurs donnent l'exemple suivant pour illustrer les déficiences des systèmes documentaires.

Soit un échantillon de trois documents,

---

<sup>39</sup> Cf. Gomez *et alii*, *JASIS* 1990: "All the right words : finding what you want as a function of indexing vocabulary".

Soit la requête "*IDF in computer-based information look-up*" (IDF est une mesure de pondération des termes de la requête dans certains modèles de recherche d'information),

Sachant que : REL signifie *relevant* c'est-à-dire "pertinent" (pour l'utilisateur qui a formulé la requête) et que MATCH signifie "tel mot est contenu dans le texte de la requête",

On obtient alors le tableau suivant :

Requête : "*IDF in computer-based information look-up*".

**Tableau 21 :**

***Illustration de la déficience des méthodes documentaires selon les auteurs de LSA.***

	Access	Docu- ment	Retrie- val	Infor- mation	Theory	Data- base	Index- ing	Com- puter	REL	MAT- CH
Doc 1	x	x	x			x	x		R	
Doc 2				x*	x			x*		M
Doc 3			x*					x*	R	M

x\* = termes apparaissant à la fois dans la requête et le document

D'après Deewester *et alii* [1990, p. 392].

Les deux documents contenant deux mots de la requête ("information" et "computer") seront retournés à l'utilisateur alors que pourtant l'un de ces deux documents n'est pas pertinent : il s'agit du document 2 traitant de la théorie de l'information. En revanche, le troisième document indexé avec les mots "access-Document-Retrieval" ne sera pas retourné à l'utilisateur alors qu'il est pertinent eu égard à la requête et cela bien qu'il n'en contienne aucun des termes.

Le bilan est donc : une réponse fausse (bruit) et une omission (silence) pour une réponse vraie.

Le problème de la synonymie reflète en fait l'opposition du point de vue de l'utilisateur et celui du système : pour l'utilisateur le terme "look-up" devrait être dans les documents tandis que dans le système ce sont les termes "access" et "retrieval" qui sont présents. La requête

comme le document sont des échantillons, des sélections de ce qui pourrait être écrit sur le sujet.

L'objectif des auteurs est alors de trouver un moyen de déduire quels termes sont réellement impliqués par une requête ou quels termes sont applicables à un document (c'est-à-dire la sémantique latente) sur la base de l'échantillon effectivement trouvé là.

#### **d) Comparaison entre AFC ET LSA**

La logique de recherche de structure de LSA étant proche de l'analyse factorielle des correspondances que nous utilisons dans nos expérimentations (Partie III), nous avons voulu comparer cette nouvelle méthode américaine à l'AFC. En effet, LSA et AFC relèvent de deux traditions de l'analyse factorielle qui s'ignorent<sup>40</sup>. Une comparaison entre les deux méthodes nous a donc semblé intéressante.

Pour réaliser cette comparaison, nous avons repris les données de Berry et DUMAIS 1994 et comparé nos résultats à ceux obtenus à partir de la SVD.

Les auteurs traitent un ensemble de titres issus de revues spécialisées en mathématiques dont la requête "theory and application", conçue comme un pseudo-document. L'objectif est de trouver le meilleur angle possible contenant le vecteur-requête pour délimiter la zone des documents pertinents à retourner à l'utilisateur.

Nous reproduisons ci-dessous le tableau des données traitées par les auteurs.

---

<sup>40</sup> Les auteurs français citent parfois la méthode américaine LSA (voir par exemple Lebart et Salem [1994] p.244; p.261) mais les auteurs de LSA ne mentionnent pas l'AFC.

**Tableau 22 :**  
**Tableau des données traitées par Berry et Dumais [1994]**

Contenu : Titres d'ouvrages contenus dans la revue SIAM.

Les termes soulignés sont ceux qui entrent dans la matrice termes\*documents à partir desquels la SVD est effectuée.

Les titres contenant les mots de la requête "theory (and) application" sont mis en caractères gras.

Les documents ajoutés pour la mise à jour sont en italique.

Le singulier et le pluriel sont comptés comme une seule forme.

CODE	TITRES
B1	A Course on Integral <u>Equation</u>
B2	Attractors for semigroups and Evolution <u>Equations</u>
<b>B3</b>	<b>Automatic Differentiation of <u>Algorithms</u> : <u>Theory</u>, <u>Implementation</u>, and</b>
B4	<b><u>Application</u></b>
B5	Geometrical Aspects of <u>Partial Differential Equations</u> Ideals, Varieties and <u>Algorithms</u> - An <u>Introduction</u> to Computational Algebraic
B6	Geometry and Commutative Algebra
B7	<u>Introduction</u> to Hamiltonian Dynamical <u>Systems</u> and the N-Body <u>Problem</u>
B8	Knapsack <u>Problems</u> : <u>Algorithms</u> and Computer <u>Implementations</u>
B9	<u>Methods</u> of Solving Singular <u>Systems</u> of <u>Ordinary Differential Equations</u>
B10	<u>Nonlinear Systems</u>
B11	<u>Ordinary Differential Equations</u>
B12	<u>Oscillation Theory</u> for Neutral <u>Differential Equations</u> with <u>Delay</u>
B13	<u>Oscillation Theory</u> of <u>Delay Differential Equations</u>
B14	Pseudodifferential Operators and <u>Nonlinear Partial Differential Equations</u>
B15	Sinc <u>Methods</u> for Quadrature and <u>Differential Equations</u>
B16	Stability of Stochastic <u>Differential Equations</u> with Respect to Semi-Martingales
<b>B17</b>	<b>The Boundary <u>Integral</u> Approach to Static and Dynamic Contact <u>Problems</u></b> <b>The Double Mellin-Barnes Type <u>Integrals</u> and Their <u>Applications</u> to Convolution</b> <b><u>Theory</u></b>
<i>B18</i>	
<i>B19</i>	<i><u>Systems of Nonlinear Equations</u></i>
<b>B20</b>	<i><u>Ordinary Algorithms for Integral and Differential Equations</u></i> <b><u>Ordinary Applications of Oscillation Theory</u></b>

Nous reproduisons figures n° 4 à 10 les résultats des traitements de ces données en reprenant les résultats de la SVD donnés par les auteurs pour les comparer à ceux de l'AFC.

Les comparaisons ont porté sur :

- la position de la requête "theory application" (il faut comprendre "theory and application") : elle est représentée sur les graphiques des figures 4 à 6. Nous précisons toutefois que le graphique correspondant à l'AFC a été réalisé avec les documents B1 à B20 alors que celui représentant la SVD a été réalisé avec les documents B1 à B17. Nous avons fait les deux graphiques : de B1 à B17 et de B1 à B20. La position des points B1 à B17 n'est en rien modifiée.
- la stabilité, au sein de chacune des méthodes, de la position des documents ajoutés pour la mise à jour par rapport aux autres, placés tantôt en éléments supplémentaires tantôt traités au même titre que les autres documents : figure 7 à 8 pour la SVD et 9 à 10 pour l'AFC.

1. Concernant la position de la requête on peut dire :

- A propos des résultats de la SVD, le seuil de valeur du cosinus choisi par les auteurs délimite un angle autour de la requête (la zone grisée sur le graphique : figure 4), qui contient 6 documents : B6, B16, B5, B7, B17, B3. Les auteurs jugent cette sélection de document satisfaisante.
- L'AFC place la requête (le point "query" sur le graphique : figure 6) très exactement entre les termes "theory" et "application", *ce qui n'est pas le cas avec la SVD*. Les documents comprenant les termes "theory" et "application" (B3, B17, B20) sont bien ceux qui sont le plus proches du point "query" sur le graphique.

Le problème théorie et application est lié d'après l'AFC à :

- "oscillation" par le document B20 "Ordinary Applications of Oscillation Theory"
  - "integral" par le document B17 "The Double Mellin-Barnes Type Integrals and their Applications to Convolutions Theory". Cela induit une proximité avec les documents B1 et B19 : "A course on Integral Equations" et "Ordinary Applications for Integral and Differential Equations"
  - "implementation" et "algorithm" par le document B3 : "Automatic Differentiation of Algorithms : Theory, Implementation, and Application".
- Dans l'AFC comme dans la SVD les documents relatifs aux équations différentielles et aux systèmes non linéaires se situent sur le graphique dans la partie opposée à celle comprenant le point requête.

2. Concernant la stabilité, au sein de chacune des méthodes, de la position des documents ajoutés pour la mise à jour (documents B18, B19, B20).

Dans le cas de la matrice termes\*documents reconstruite par SVD (figure 8 ), comme dans le cas de l'AFC traitant les documents issus de la mise à jour en tant qu'éléments supplémentaires (figure 10 ), le terme "ordinary" migre.

Quatre documents contiennent "Ordinary" : B8, B10, B19, B20. Seul B19 est lié à la théorie de l'oscillation ; les 3 autres sont liés au thème des équations différentielles. Dans les analyses non mises à jour, dans l'AFC traitant les documents ajoutés au même titre que les autres et dans l'une des méthodes de mise à jour de SVD "ordinary" se trouve proche de "differential" et "equations".

Dans le second cas (AFC) il se trouve à "mi-chemin" de "oscillation" et "differential", "equations".

Dans le cas de la SVD sa position est difficilement interprétable.

En conclusion de cette comparaison, il nous semble que la représentation géométrique issue de l'AFC donne des résultats beaucoup plus "lisibles" et simples à interpréter. Il est vrai que les auteurs de la LSA utilisent ce type de méthode pour déterminer une zone de documents pouvant être renvoyés à l'utilisateur (zone grisée du graphique p.109) sans rechercher une interprétation générale de l'espace mots\*documents obtenu. Or cette zone nous semble un peu trop large pour être réellement pertinente. L'interprétation des nuages de points résultant du traitement par AFC nous semble plus intéressante.

Figure n°4

ANALYSE PAR SVD D'APRES BERRY et DUMAIS [1994 ](figure 4)  
et position d'un requête (QUERY) qui est "application theory"

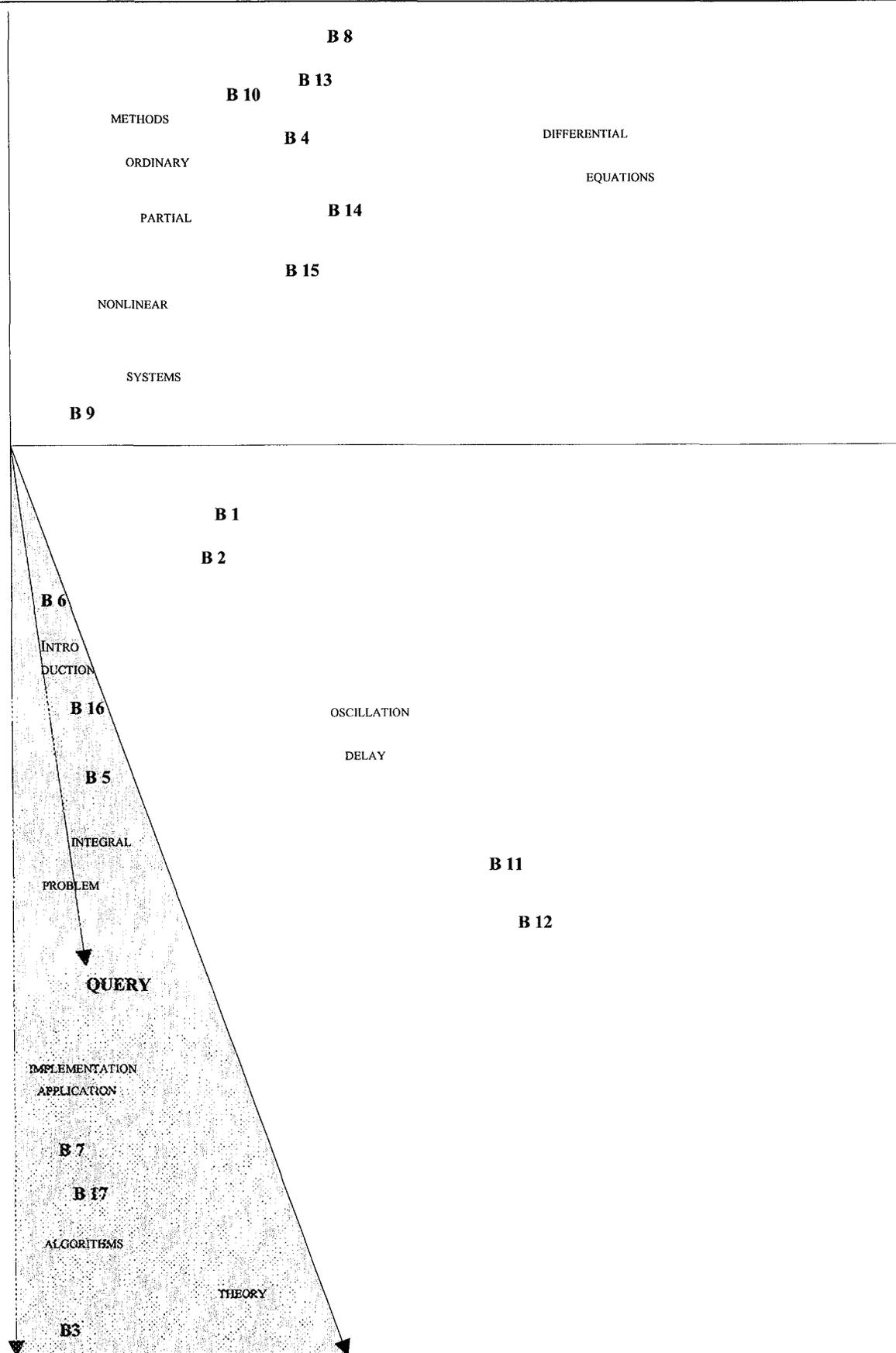


Figure n°5

ANALYSE PAR AFC DU MEME TABLEAU DE DONNEES QUE DUMAIS ET BERRY 1994 SANS POSITIONNER LA REQUETE

x = 1 ( 21.00%), y = 2 ( 16.57%)

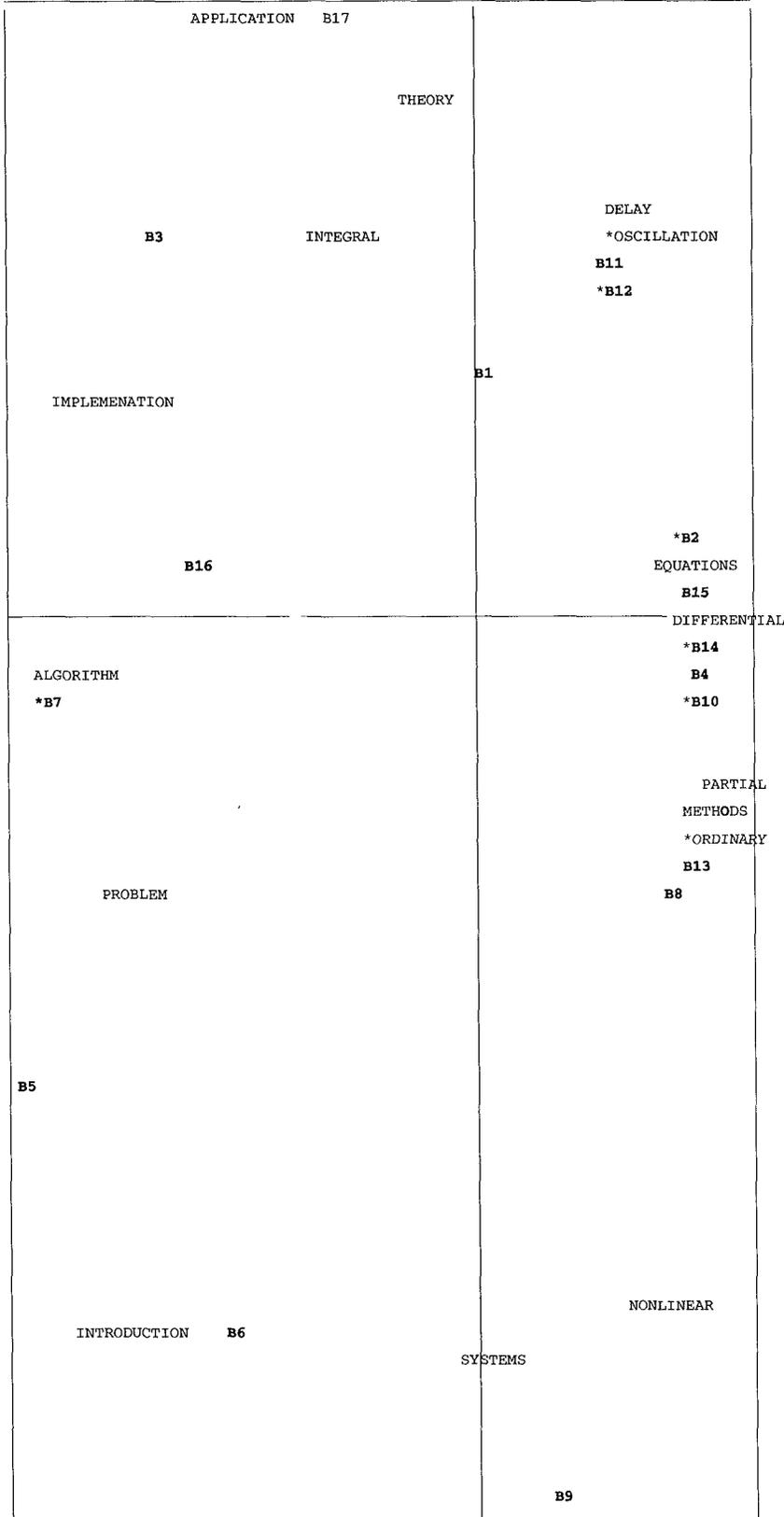
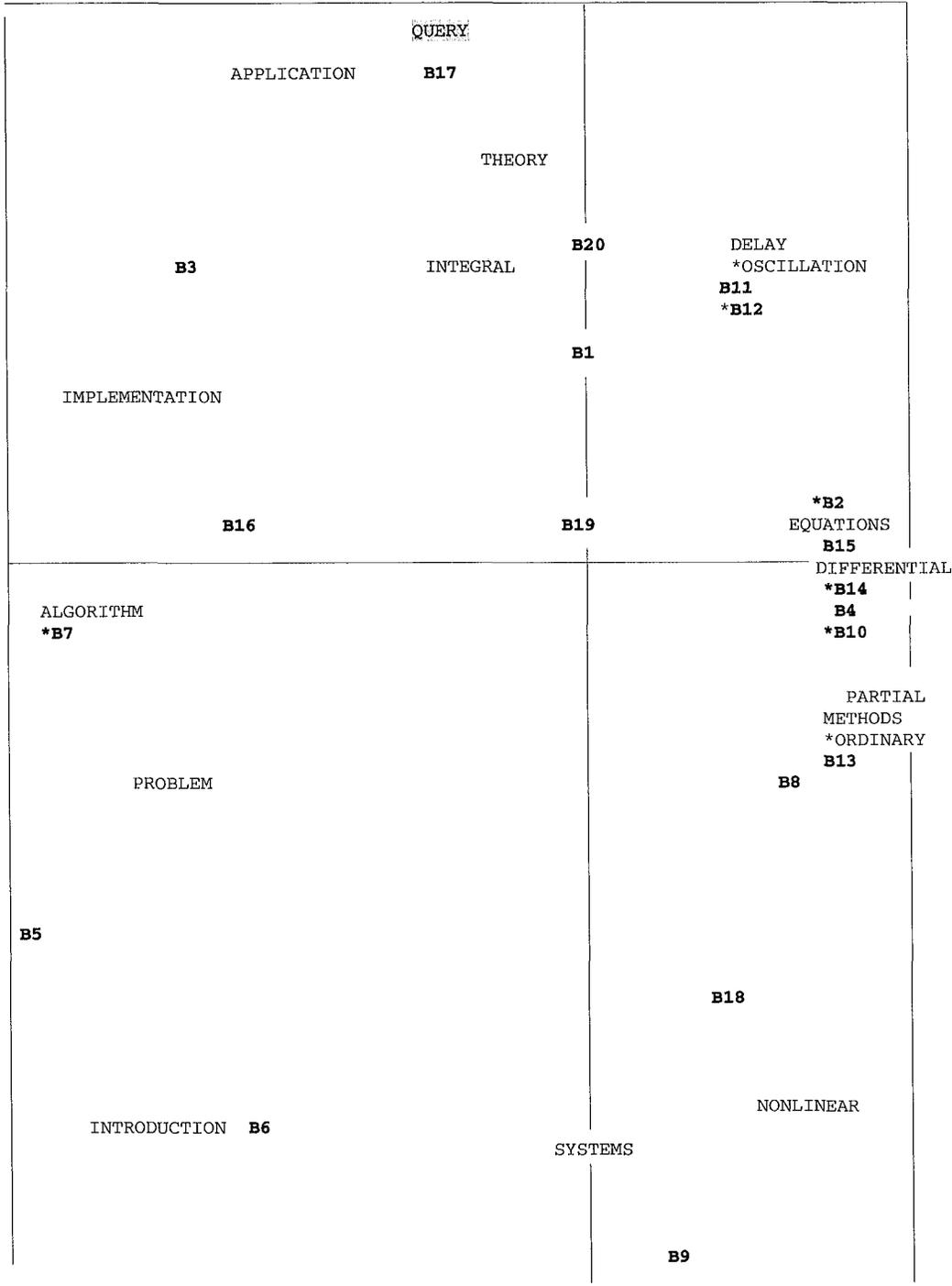


Figure n°6

Position de la requête (point *QUERY*) "application theory" avec le traitement des données par AFC

x = 1 ( 21.00%), y = 2 ( 16.57%)



positions des points superposés, nø carac., nø ligne

- \*ORDI:(c 72, 1 33)      \*OSCI:(c 64, 1 11)      \*B2:(c 71, 1 24)
- \*B7:(c 3, 1 28)      \*B10:(c 72, 1 28)      \*B12:(c 63, 1 13)
- \*B14:(c 72, 1 28)

Figure n°7

ANALYSE PAR SVD DU MEME TABLEAU EN AJOUTANT  
LES DOCUMENTS B18 B19 B20 EN TANT QUE VARIABLES ILLUSTRATIVES  
(BERRY DUMAIS 1994 fig 7)

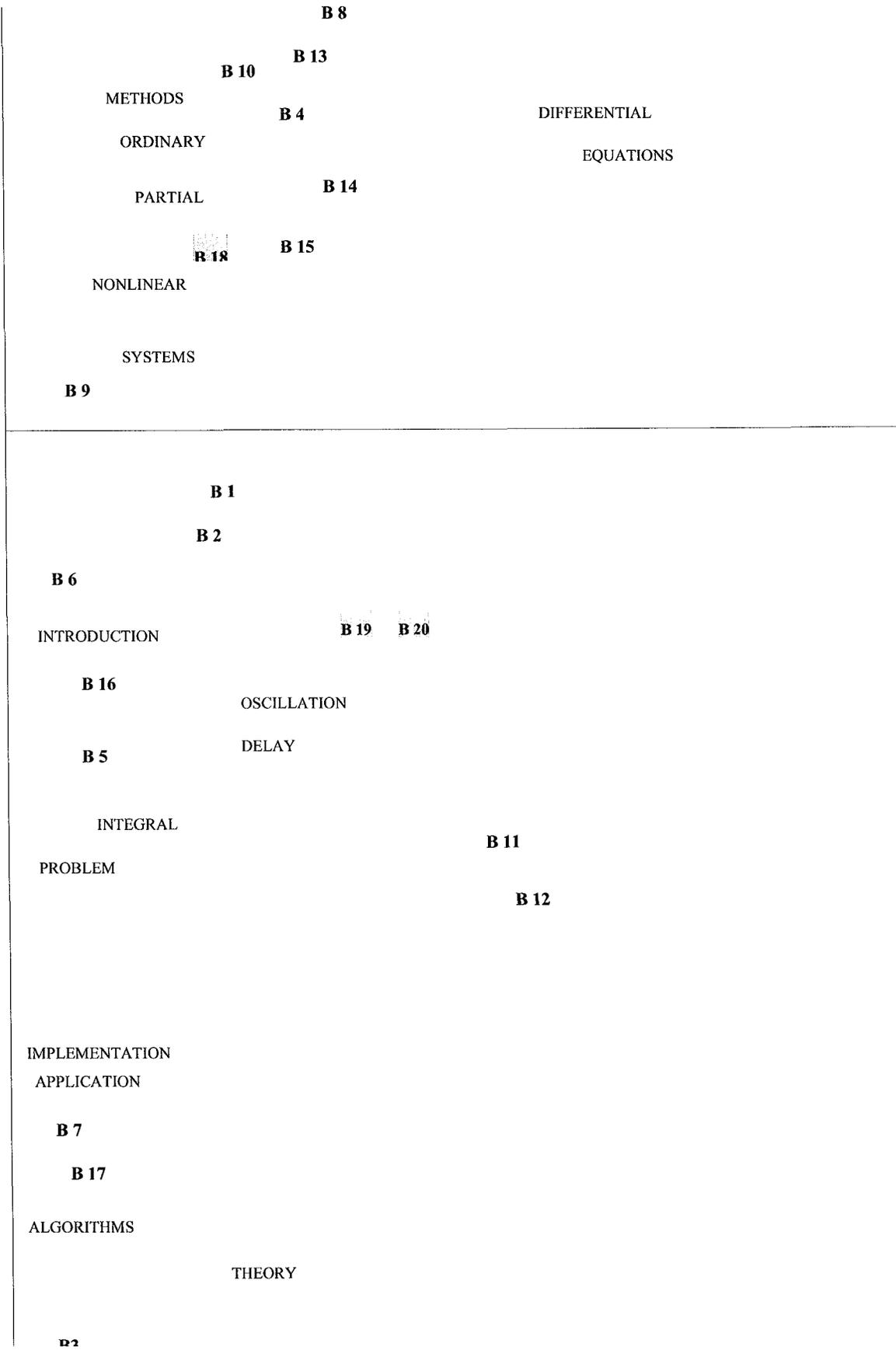


Figure n°8

ANALYSE PAR SVD D'APRES BERRY et DUMAIS 1994 (Figure 8 ) EN AJOUTANT  
LES DOCUMENTS B18 B19 B20/ "using the SVD of a reconstructed term-document matrix"

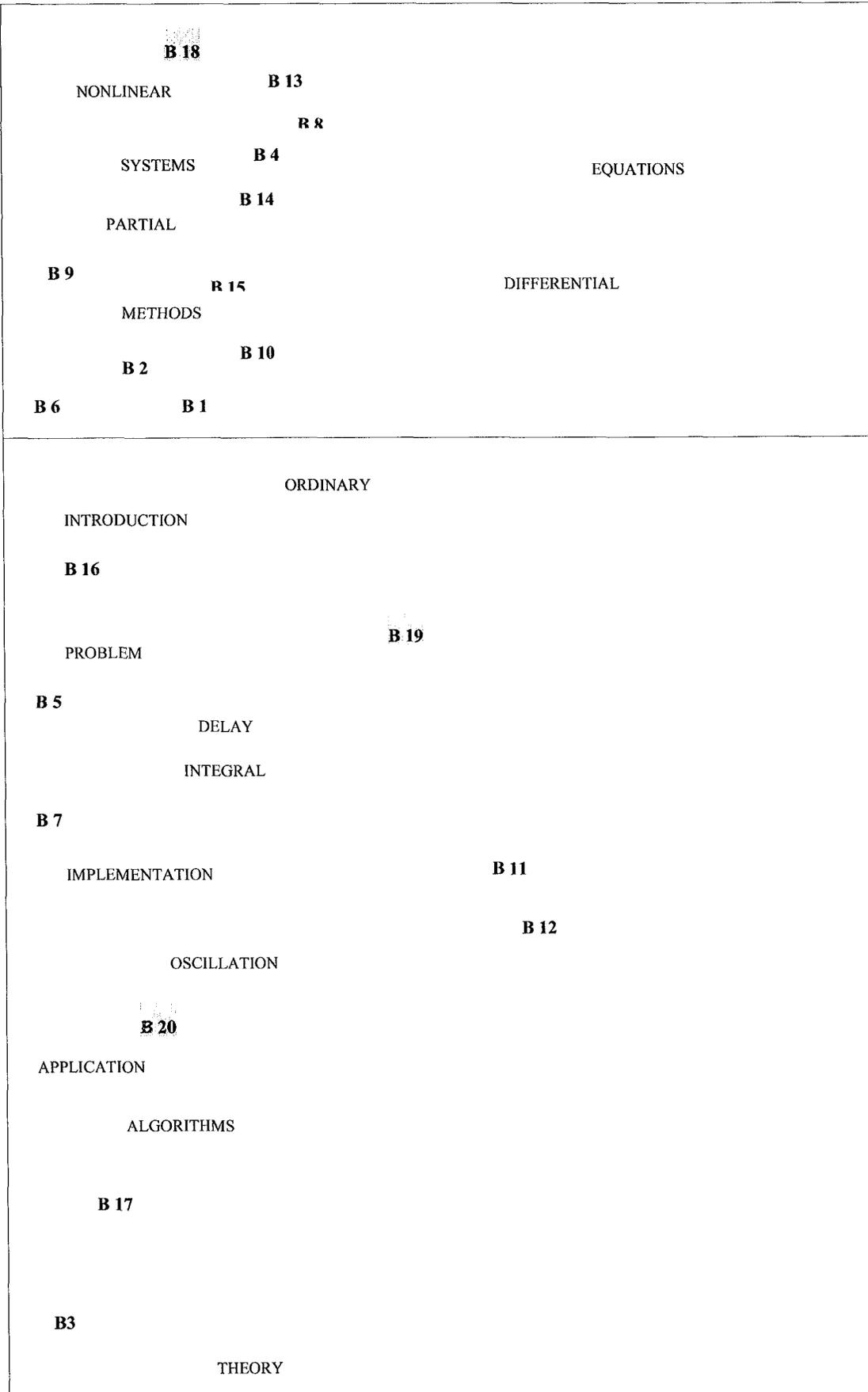
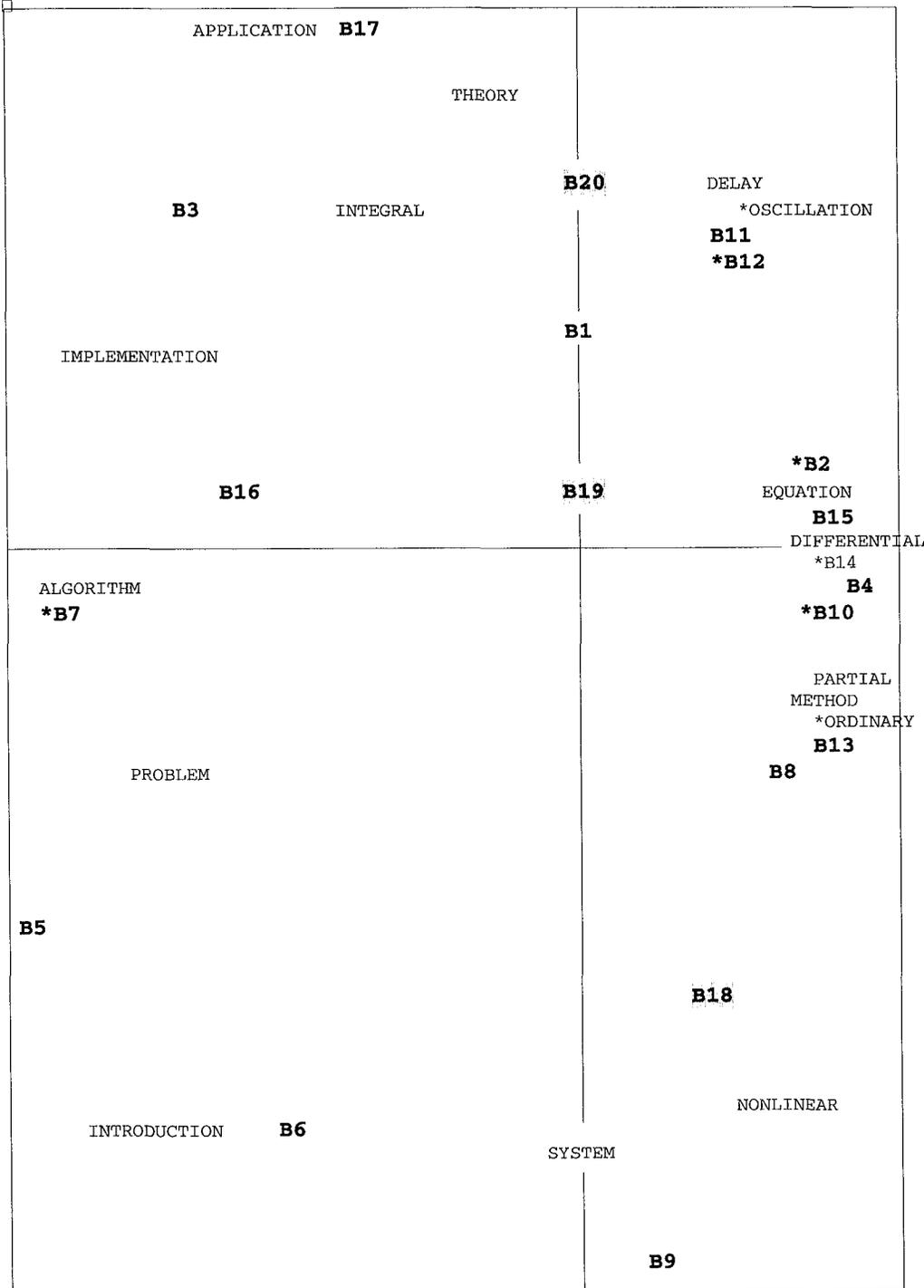


Figure n°9

**AFC SUR LES DONNEES DE BERRY DUMAIS 1994 EN AJOUTANT LES DOCUMENTS B18 B19 B20 EN TANT QUE VARIABLES ILLUSTRATIVES**

x = 1 ( 21.00%), y = 2 ( 16.57%)



positions des points superpos,s, n° carac., n° ligne

\*ORDI:(c 72, l 30)  
 \*B7:(c 3, l 25)  
 \*B14:(c 72, l 25)

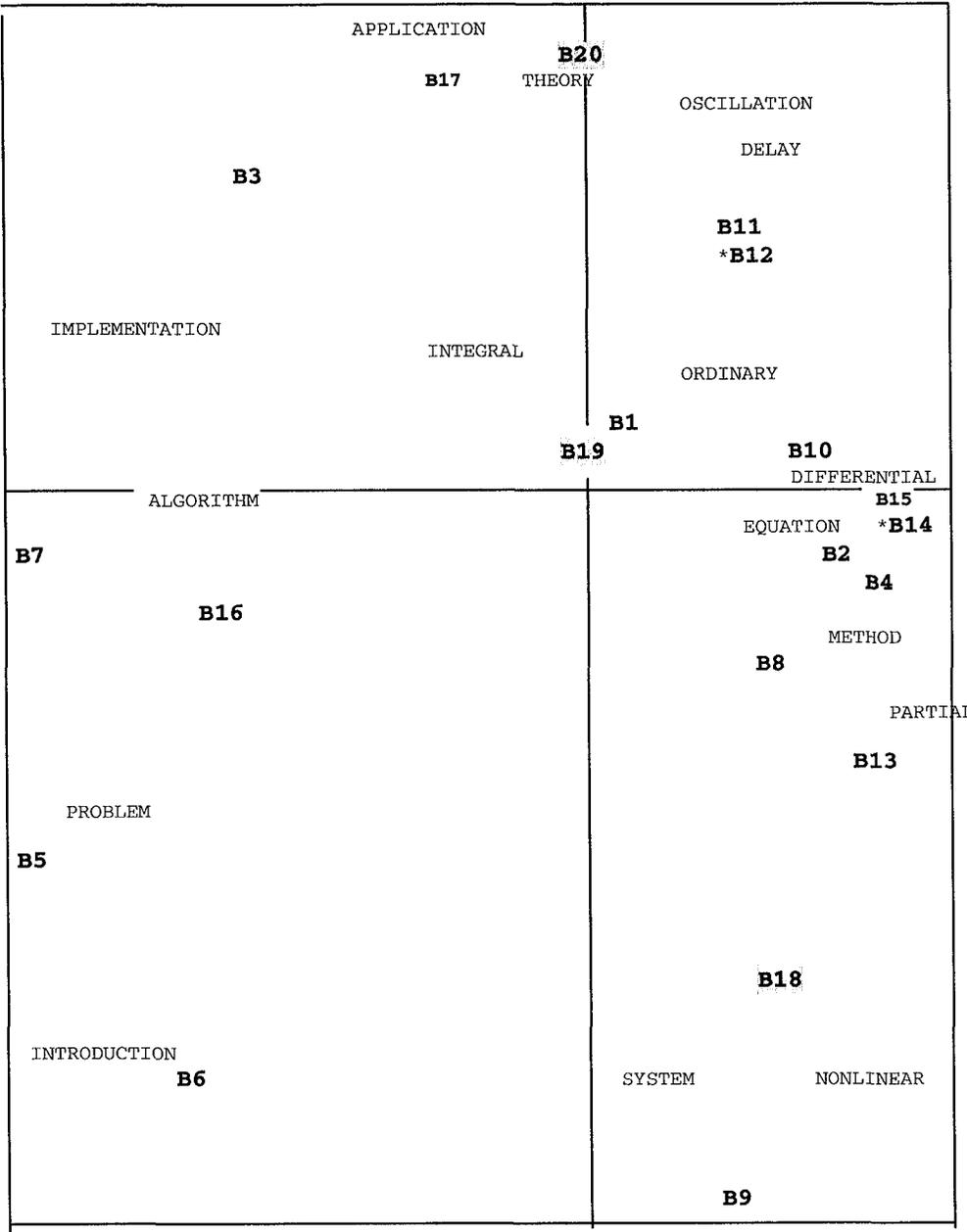
\*OSCI:(c 64, l 8)  
 \*B10:(c 72, l 25)

\*B2:(c 71, l 21)  
 \*B12:(c 63, l 10)

**Figure 10**

**AFC SUR LES DONNEES DE BERRY DUMAIS 1994 EN AJOUTANT LES DOCUMENTS B18 B19 B20 EN TANT QUE VARIABLES ACTIVES**

$x = 1$  ( 19.48%),  $y = 2$  ( 17.36%)



## ***II-5 CONCLUSION***

Nous venons de voir en quel sens l'intelligence économique peut solliciter le traitement des documents.

L'essentiel du traitement du contenu des documents repose sur la mise en contexte des termes et la mise en relation des contextes plutôt que des mots. Les différences d'approche reposent sur la définition du contexte:

- logique : syntagme nominal ou contexte syntaxiquement construit.
- sémantique : arborescence et relation transversale (l'antonymie est transversale et non hiérarchique). Ici le contexte distingue la nature des relations sémantiques liant les termes entre eux.
- statistique : cooccurrence c'est-à-dire co-apparition des mots dans une même unité et relation entre ces cooccurrences dans des agrégats (mots associés), ou bien profils d'utilisation des mots dans le corpus et profils du vocabulaire des textes du corpus mis en relation par un calcul de distance et représenté simultanément (Analyse factorielle).

Nous allons voir (Partie III) quels résultats peuvent être obtenus à partir de cette dernière notion de contexte dans le cas du traitement de larges corpus.

**Partie III**  
**PROPOSITION POUR UNE METHODE D'ANALYSE DE**  
**L'INFORMATION TEXTUELLE DESTINEE A L'INTELLIGENCE**  
**ECONOMIQUE : EXPERIMENTATIONS**

<b>PARTIE III- PROPOSITION POUR UNE METHODE D'ANALYSE DE L'INFORMATION TEXTUELLE DESTINEE A L'INTELLIGENCE ECONOMIQUE : EXPERIMENTATIONS</b>
--

### **III-1 MATERIEL ET METHODE**

#### III-1-1 LE CHOIX DE L'ANALYSE FACTORIELLE DES CORRESPONDANCES (AFC) POUR TRAITER LE TEXTE LIBRE

##### **III-1-1. 2 La démarche**

Nous avons adopté la démarche suivante :

- 1- constitution d'une base de documents sur un sujet.
- 2- Recherche d'une typologie thématique de ce sujet à partir du traitement par AFC des textes des documents. Cette notion de typologie thématique signifiait pour nous : comment obtenir une image du contenu d'une grande quantité de documents afin d'obtenir une sorte de guide de lecture de cette base ? D'autre part, dans la mesure où notre objectif est de connaître un sujet en étant le moins possible limité ni par l'*a priori* d'une requête trop précise ni par le nombre de documents contenant des réponses potentiellement pertinentes, quelle est la valeur de l'information obtenue par la méthode proposée ? Cette valeur nous donnera un indicateur du caractère éventuellement stratégique de l'information obtenue et de son utilisation possible pour l'intelligence économique.

Nous n'utilisons donc pas l'AFC pour sélectionner des documents en fonction d'une requête contrairement à l'utilisation américaine de la LSA. Cela est bien sûr tout à fait envisageable (voir II-4-3-1, 2 : "Comparaison entre LSA et AFC" p .105) en traitant la requête comme un pseudo-document, mais nous ne cherchons n'est pas à construire un moteur de recherche.

Cette démarche est particulière en ce sens qu'elle est intermédiaire entre une méthodologie documentaire qui vise à sélectionner des documents à l'intérieur d'une collection en fonction d'une requête et une démarche strictement bibliométrique dont l'objectif est d'obtenir des informations de synthèse sur les champs secondaires (non textuels) des documents : *ici nous cherchons à la fois à répondre à un besoin d'information (démarche documentaire classique) et à obtenir des informations de synthèse sur l'ensemble de la base.*

Nous avons pu constater que cette approche était souvent incomprise de la part des spécialistes de l'informatique documentaire. L'objection consiste à faire remarquer que la base initiale étant sélectionnée, pour la part des informations issues de l'interrogation des serveurs, selon les méthodes habituelles, c'est-à-dire par mots-clés et équations logiques, on ne voit pas comment il est possible de ne pas introduire d'*a priori* dans la recherche.

En réalité l'*a priori* se réduit à l'expression "minimale" du sujet en termes de précision : une expression voire une équation logique entre deux chaînes de caractères simples, bref le point départ permettant de constituer une base de documents *spécialisée* ou ce que l'on peut appeler un corpus.

Notre problème est celui d'un traitement de l'information destiné à l'intelligence économique : il ne s'agit pas pour nous de "seulement" sélectionner des documents, encore moins de le faire à partir d'une collection hétérogène quant à son contenu comme peut l'être celui d'une bibliothèque ou d'un serveur commercial. Sur un serveur commercial on trouvera des bases de données spécialisées par discipline (par exemple RAPRA : polymères, matériaux et procédés de transformation) ou par champ (FRANCIS : base de l'INIST<sup>41</sup> en sciences de l'homme et de la société).

L'intelligence économique travaille sur des sujets nettement plus délimités que ceux que nous venons d'évoquer (disciplines ou champs généraux) et en même temps beaucoup moins précis qu'une recherche de données<sup>42</sup> : les méthodes documentaires classiques ne lui suffisent pas. La sélection des documents n'est pas un objectif en soi : ***le référentiel de la sélection (le contenu du corpus dans son ensemble) est aussi important que la sélection elle-même (le contenu de quelques documents se rapportant précisément à la requête)***. L'importance de la typologie thématique porte précisément sur ce point : elle permet d'obtenir à la fois le contenu de la base pour avoir une image du sujet et de sélectionner des documents à partir de cette image du sujet (ce que nous appelons le référentiel). Mais la seule représentation du sujet dans sa totalité, sans objectif de sélection, peut être recherchée pour elle-même. En d'autres termes, notre démarche se différencie de celle d'un moteur de recherche dans la mesure où notre résultat n'est pas seulement une sélection de documents mais un ensemble de thèmes illustrés par des documents.

---

<sup>41</sup> Institut National de l'Information Scientifique et Technique

<sup>42</sup> Voir la distinction entre les différents besoins d'information ci-dessus p.73. La recherche de données est le premier niveau de ces besoins.

### III-1-1. 3 Le problème méthodologique

Pour ce faire nous avons voulu tester la validité d'un traitement massif sur les résumés de plusieurs centaines voire de milliers de documents ou de brèves de presse sans lemmatisation des termes. La difficulté de cette approche est que les matrices mots\*documents ainsi constituées à partir de la quasi-totalité des mots non lemmatisés comportent un très grand nombre de zéros. On parle alors de "matrices creuses". Ces matrices creuses comprennent, en moyenne, dans le cas des études réalisées par nous, environ 96% de valeurs nulles.

De plus, la taille des matrices issues du traitement des données linguistiques est importante : plusieurs centaines ou plusieurs milliers de documents sont mis en correspondance avec plusieurs centaines de mots. La variance est dispersée sur de nombreux facteurs, très loin de se résumer aux 70% habituellement représentés par les 5 premiers facteurs. Lebart, Morineau et Fénelon [1979, p.295-298] ont montré une corrélation entre le pourcentage de variance expliquée par les cinq premières valeurs propres et le nombre de variables et d'observations mis en correspondance. La chute du pourcentage d'inertie est proportionnelle tant au nombre d'observations qu'au nombre de variables<sup>43</sup>.

Dans le cadre de notre travail, la distribution des valeurs propres est telle qu'une cinquantaine de facteurs est nécessaire pour expliquer 50% de la variance.

Si l'on ajoute au cas des matrices creuses celui de la dispersion de l'information liée au traitement de grandes matrices, que peut-on dire de la cohérence sémantique des associations de mots et de documents obtenues sur chaque facteur ?

C'est ce que nous avons voulu savoir en testant l'intérêt des résultats issus du traitement de telles matrices. Si notre démarche visant à trouver une méthode de traitement de l'information documentaire appropriée à l'objectif "intelligence économique" nous a conduit à utiliser l'AFC sur les textes des documents, il nous a fallu encore évaluer la qualité de la méthode utilisée.

---

<sup>43</sup> Voir dans *Annexe III-A* p.208 les courbes d'évolution des pourcentages de variance relatifs aux 5 premières valeurs propres en fonction soit du nombre des observations, soit de celui des variables.

Nos expérimentations répondent à cet objectif. Nous en donnons la liste dans l'Annexe III-C p.210.

- Elles ont porté sur 7 jeux de données :

- Capteurs (Sensor) ; trois jeux de données traités base par base : Capteurs RAPRA 1995 ; Capteurs ISI 1995 et Capteurs mise à jour.
- Polymères à cristaux liquides (Liquid Crystal Polymer/LCP) : un jeu de données "monobase"<sup>44</sup>.
- Intelligence économique (*Competitive Intelligence*) : un jeu de données multibases.
- Le recyclage du polyuréthane : un jeu de données multibases.
- Journal Bref Rhône-Alpes : un jeu de données monobase.

- 21 Analyses Factorielles des Correspondances ont été réalisées.

### III-1-1. 4 Les motivations de l'utilisation de la méthode

Les raisons de ce choix ont déjà été soulignées tout au long de l'exposé précédent (Partie II).

#### **Nous les récapitulons :**

La majorité des ambiguïtés linguistiques étant liée à la polysémie ou à l'homographie des termes, on peut penser, comme le soulignaient Deerwester et les auteurs de la LSA [Deerwester *et alii*, 1990], qu'elles ne seront pas un obstacle déterminant avec l'analyse factorielle grâce à la prise en compte du profil de l'utilisation d'un mot dans l'ensemble des documents

2- Outre le comportement favorable de l'AFC à l'égard des ambiguïtés linguistiques, l'avantage de l'utilisation de cette méthode sur les données linguistiques, est qu'elle ne fait pas appel à des ressources externes. Dans la mesure où l'objectif est de construire un système documentaire dédié à l'intelligence économique cet aspect est important.

3- L'AFC est une méthode permettant le classement simultané des mots et des documents.

### III-1-2 LES TEXTES TRAITES

Ces textes sont des résumés d'articles ou des brèves de presse économique.

Une seule base de documents est en français ; les autres sont en anglais.

Pour les disciplines concernées les serveurs suivants ont été utilisés :

- Sciences et techniques (articles et brevets) : RAPRA, ISI.
- Entreprises : ABI INFORM, PROMT, TEXTLINE GLOBAL NEWS.
- Brèves de presse : BREF RHONE-ALPES
- La base "Recyclage du Polyuréthane" comprend : DERWENT (brevets), CHEMICAL ABSTRACTS, ENERGY SCIENCE, et TECHNOLOGY ENVIROLINE.

Taille des bases traitées :

- de 386 documents à 987 documents soit des bases de 0.5 à 2.7 Mo (Megaoctets).
- De 300 à 933 mots entrant dans les différentes analyses.

### III-1-3 METHODE

A partir du traitement par Analyse Factorielle des Correspondances d'une table de contingence (ici un tableau mots\*documents), plusieurs séries de paramètres sont extraites :

- les valeurs propres et les pourcentages d'inertie représentés par chaque axe (les facteurs). Ces valeurs mesurent la "part d'information" représentée par chaque axe.
- les coordonnées de chaque point sur les axes.
- les contributions qui décrivent la part prise par un élément (ligne ou colonne) dans la construction d'un axe factoriel.
- les cosinus carrés qui mesurent la qualité de représentation de chaque élément sur les axes.

Nous avons considéré que l'interprétation des axes devait s'opérer sur la valeur de contribution des mots et sur celle de la qualité de représentation des documents.

---

<sup>44</sup> "monobase" c'est-à-dire provenant d'une seule et même origine.

Afin de rendre l'interprétation possible, nous avons édité exclusivement les mots contribuant le plus à la construction des facteurs (seuil  $\geq 2$  fois à la moyenne des contributions pour chacun des axes) et les documents les mieux représentés sur ces facteurs (seuil choisi de 67 pour mille<sup>45</sup>)

A partir de ce critère général, nous avons fait varier les seuils et les manières de procéder (voir chapitre III-2. 2 p.145-153) afin de savoir à quelles conditions les associations de mots et de documents avaient une "valeur sémantique" et de quelle manière cette dernière variait.

### III-1-4 LOGICIEL UTILISE

Pour la réalisation de l'ensemble des travaux dont nous présentons ici les résultats, nous avons utilisé le logiciel de calcul d'une équipe Inserm<sup>46</sup> : le logiciel BI (BI@logInserm 1979, 1987, 1993). Ce logiciel permet le traitement des données numériques et alpha-numériques c'est-à-dire à la fois :

- la saisie
- la gestion
- l'édition
- les statistiques
- l'analyse des données

Plus précisément le logiciel BI permet :

- 1 - la gestion simple de données numériques et alpha-numériques,
- 2 - le codage de ces données,
- 3 - la mise en œuvre d'outils statistiques
- 4 - la préparation des données en vue de traitement par d'autres logiciels, en particulier les logiciels d'analyse de données de l'ADDAD (Association pour la Développement et la Diffusion de l'Analyse des Données).

---

<sup>45</sup> Sachant que le cosinus carré est égal à 1 et est multiplié par mille pour une meilleure lisibilité.

<sup>46</sup> Il s'agit de l'équipe INSERM du Département de Santé Publique de la Faculté de Médecine de Rennes.

Les objets manipulés par BI sont composés d'une liste de valeurs alpha-numériques, suivie d'une liste de valeurs numériques. La partie représentant la liste alpha-numérique est souvent appelée REFERENCE de l'objet. La liste numérique est souvent considérée comme un VECTEUR à  $p$  dimensions et contient les valeurs prises par l'objet pour ces  $p$  descripteurs.

Pour réaliser un travail avec BI, l'utilisateur dispose d'une suite de commandes (ou primitives) qui effectuent des tâches précises. La réalisation d'un calcul complet est souvent l'aboutissement de la mise en œuvre de plusieurs commandes subséquentes. Chaque commande de BI traite le fichier courant en entrée et génère un nouveau fichier après exécution, ce nouveau fichier devenant automatiquement le nouveau fichier courant. Entre chaque commande, il est possible de sauvegarder le fichier courant ou de provoquer son changement. Le logiciel peut également s'exécuter en mode fichier de commande. Ce type de procédure est particulièrement utile pour les travaux répétitifs ou pour le traitement de grands fichiers.

L'accès aux commandes se fait par trois voies :

- 1 - en fournissant le numéro de la commande
- 2 - en donnant le nom du module
- 3 - en donnant une chaîne de caractères contenue dans le descriptif de la commande

Une fonction AIDE en ligne permet de connaître la liste des commandes existantes. Il suffit de taper "?" à l'invite du programme.

### III-1-5 PRESENTATION DES RESULTATS

Les sujets et les différentes analyses effectuées sont numérotées dans l'Annexe III-C p.210.

Les résultats cités sont, soit insérés dans le corps du texte de notre étude, soit présentés à part dans la partie "Résultats Détaillés".

La table suivante restitue l'ensemble des résultats cités dans cette étude. Ils sont numérotés de 1 à 18 et localisés par le numéro de page.

## Table des résultats cités

Numéro et titre	Page
1. Qu'est-ce qu'une métaclé ?	127
2. Le ciblage des données marché	129-130; 176-178
3. Le thème "Formation" dans l'analyse du journal Bref Rhône-Alpes	132-134
4. Exemple de typologie thématique	179-192
5. Exemple d'associations "triviales"	136-137
6. Exemple de difficultés d'interprétation à partir des mots de la métaclé	138-140
7. La notion d'hyperplan d'un document : exemple.	141-142
8. Exemple de cartographie thématique	144
9. Comparaison entre trois manières de réaliser l'AFC à partir des données "Polymères à cristaux liquides"	148
10. Comparaison entre l'analyse 746 et l'analyse 673 : document spéci-fiques de l'analyse 746	150
11. Comparaison entre l'analyse 746 et l'analyse 673 : documents spéci-fiques de l'analyse 673	151
12. Analyse des documents rejetés par une analyse précédente : spécificité du thème "structure du LCP"	193-195
13. Analyse des documents rejetés par une analyse précédente : spécificité du thème "mélange à base de LCP"	196
14. Evolution chronologique du vocabulaire	159; 197-198
15. Désambiguïsation sans dictionnaire : cas de l'analyse Capteurs	199-207
16. Présence du mot "membrane" au singulier seulement	161-162
17. Sens de la métaclé avec "membrane" au singulier seulement	163
18. Association sur la même métaclé du singulier et du pluriel de la forme "membrane"	164-165

## **III-2 RESULTATS**

### **INTRODUCTION**

Le très grand nombre d'analyses réalisées durant cette étude nous a permis de mettre en évidence quelques caractéristiques importantes liées à l'utilisation de l'Analyse Factorielle des Correspondances en documentation, non plus seulement pour en montrer l'intérêt sur un jeu de données à partir de deux facteurs particulièrement intéressants, mais bien pour en évaluer les potentialités (forces et faiblesses) à partir du problème méthodologique exposé ci-dessus (voir p. 118-119). En d'autres termes, le traitement des documents par AFC n'est pas utilisé ici à titre illustratif mais bien pour être évalué en fonction d'un objectif ("l'intelligence économique").

Cinq caractéristiques sont examinées :

1. la cohérence des associations
2. la variation de la typologie thématique obtenue par AFC sur une même base de documents selon que l'on choisit tel ou tel principe de sélection des mots
3. la perte d'information et plus précisément, à travers l'analyse des documents rejetés par une première analyse factorielle, l'examen d'une méthodologie visant à obtenir une analyse de contenu la plus exhaustive possible, sans toutefois céder à la surinformation.
4. La différence entre une AFC et un état de l'art réalisé par un groupe de travail.
5. La possibilité de mettre en évidence une analyse chronologique des thèmes.

#### **III-2-1 LA COHERENCE DES ASSOCIATIONS**

##### **III-2-1.1 Interprétation par "métaclé"**

Notre interprétation des analyses repose sur une double représentation des données : l'interprétation des associations par pôle positif ou négatif de chaque facteur et celle que peuvent apporter les cartes factorielles mettant en relation chaque pôle d'association.

Nous appellerons "métaclé" les pôles positifs ou négatifs de chaque facteur. La métaclé représentait, pour l'équipe de travail avec laquelle nous avons effectué ces travaux (Michel Kerbaol), le niveau d'intégration supérieur au mot-clé : un ensemble de mots-clés décrivant un thème et qui pouvait ensuite être réutilisé pour sélectionner une sous-base à l'intérieur du

corpus de départ. Le mot-clé représente le contenu d'un texte; la métaclé caractérise un sous-ensemble de textes.

Nous avons souhaité, pour notre part, ne pas limiter la métaclé aux mots-clés mais l'étendre aux textes associés (c'est-à-dire aux documents) afin de savoir s'il était possible d'interpréter à la fois l'association de mots et l'association de textes *en se servant de la première comme d'un guide de lecture pour la seconde*. C'est en ce sens que nous avons voulu évaluer si l'analyse est cohérente, c'est-à-dire si chaque métaclé représente un thème.

Notre démarche générale est donc la suivante :

-on constitue une base de documents sur un sujet donné

-on prépare les données : dédoublonnage (il s'agit d'enlever les doublets et, d'une façon générale, tous les documents identiques se trouvant en plusieurs exemplaires), reformatage (avec un éditeur de textes) des documents afin que ceux-ci soient tous au même format.

-on construit le dictionnaire des mots des textes de la base et les index correspondants.

-on constitue les tableaux de correspondance et on réalise les calculs.

-on interprète chaque métaclé afin de juger de la cohérence de chaque association de mots et de documents en les considérant une à une. L'ensemble des métaclés nous donne les thèmes de la base de documents

-on envisage l'ensemble des métaclés et l'on juge de la typologie thématique obtenue.

## RESULTATS N°1

### Qu'est-ce-qu'une métaclé ?

#### Un exemple

#### RESULTATS DU TRAITEMENT PAR AFC DE LA BASE IEC.

Cette base concerne le sujet "competitive intelligence"

**Il y a 10 mots explicatifs.**

#### **POUR LE FACTEUR 3 LISTE DES MOTS A COORDONNEE NEGATIVE.**

SEUILS POUR LA CONTRIBUTION 2\*MOY = 7.84 3\*MOY = 11.76

LIGNE	FACTEUR	QUALITE	CONTRI	*(voir ci-dessous la signification des colonnes)	
31	-0.68	135.37	41.96	118	INTELLIGENCE
34	-0.94	42.21	7.96	115	INDUSTRIAL
35	-1.01	40.72	23.55	156	NOTES
36	-1.06	48.32	8.56	96	GOVERNMENT
37	-1.15	84.43	26.49	210	SHOULD
38	-1.31	92.90	23.34	70	ECONOMIC
40	-1.46	118.97	31.71	205	SECURITY
44	-1.82	133.38	35.31	88	FOREIGN
46	-1.95	139.21	27.49	246	WAR
50	-2.31	159.04	61.17	77	ESPIONAGE

- **Qualité=67 (10 ) documents sélectionnés**

Exemple de l'un des dix documents :

@AB The only way for many companies to grow today is at the expense of their competition. If they do not, competitors will carve away their market share. A company can get a jump on the competition by knowing what the competition is up to and staying one step ahead - a process known in business as competitive **intelligence**. More executives are realizing the need to conduct competitive intelligence operations. Business **intelligence** may get confused with business **espionage**, which is illegal. When business **intelligence** operations are run successfully, the return on investment can be enormous. According to Robert E. Flynn, chairman and CEO of NutraSweet Co., competitive **intelligence** is worth up to \$50 million a year to the company.

#### *Interprétation de la métaclé*

L'association des dix mots sur le pôle négatif de ce facteur 3 renvoie au thème de la comparaison entre intelligence économique et espionnage économique. Les dix textes sélectionnés avec ces dix mots confirment ce thème car ils sont tous centrés sur cette comparaison.

C'est une telle association de mots et de documents sur le pôle de l'un des facteurs que nous appelons une "métaclé".

*\*Signification des colonnes :*

LIGNE : numéro de la ligne sur le listing de calcul statistique

**FACTEUR : coordonnée du point sur le facteur**

**QUALITE : qualité de représentation**

**CONTRI : contribution à la valeur propre du point ( "point-mot")**

**COLONNE DE CHIFFRE : numéro du point ("point-mot")**

- l'expert se positionne par rapport à cette typologie thématique : il fait le point sur ce qui le surprend ou le conforte dans la représentation qu'il a *a priori* du sujet ; il peut reposer sa question si nécessaire ou engager une nouvelle recherche sur l'un des thèmes mis en évidence par l'analyse.

L'important est pour nous, comme nous l'avons déjà noté, que le demandeur, c'est-à-dire l'utilisateur d'un système d'intelligence économique puisse ne pas se laisser guider par ses propres *a priori* sur une question mais être à même de découvrir des thématiques inenvisagées par lui au moment de la demande d'étude.

Nous montrerons par la suite <sup>47</sup> comment nous concevons la complémentarité entre cette typologie thématique et un système documentaire classique d'interrogation par requête logique. Les mots de la métaclé peuvent alors servir alors de point de départ pour une interrogation de la base analysée.

La métaclé nous permet de passer des mots (l'ensemble des mots de la base) aux concepts (abstraction née de l'association). Elle est d'abord un moyen de laisser le contenu d'un sujet se dévoiler par le seul jeu des associations sans recours à des ressources externes (dictionnaire, thesaurus etc.). Elle est également un moyen de révéler des informations sur des "abstractions".

Ainsi pourra-t-on répondre à des besoins d'information du type de ceux que nous avons déjà mentionnés précédemment (II-3-2 : p.73) : on peut mettre en évidence des informations sur le marché des polymères à cristaux liquides par exemple. Or, sans ce type de méthode, pour sélectionner ce type d'informations, il n'y a pas d'autres recours que d'utiliser des bases économiques (études de marché) ou, si l'on cherche dans des bases autres, de choisir dans un thesaurus un réseau de termes appropriés pour décrire la notion de "marché" car aucun descripteur ne peut à lui seul résumer cette notion.

Dans l'exemple suivant nous montrons la métaclé regroupant les informations technico-économiques sur le marché des polymères à cristaux liquides.

L'interprétation repose sur la cohérence de l'association des documents. La question à laquelle la démarche interprétative répond en premier lieu est : qu'est-ce que ces documents ont de

---

<sup>47</sup> Voir III-4 : "*Quel système documentaire pour l'intelligence économique*"

commun ? Pour y répondre, les mots de la métaclé nous donnent un guide de lecture. Dans l'exemple qui suit les mots de la métaclé relève de quatre registres :

1. les noms commerciaux de matière : *Vectra, Xydar*
2. les noms de société : *Hoechst Celanese, Dartco*, noms signalant qu'il s'agit de données relatives au marché : *company, applications, products, market, plant, Japan* .
3. le vocabulaire particulier des articles commerciaux, lesquels contiennent presque toujours la phrase "*Brief details are given*" pour signaler que des détails sur les produits signalés sont mentionnés dans l'article intégral.
4. Associés à cette dernière catégorie des mots tels que *resistance, plastics* pour désigner les matériaux.

## RESULTATS N°2

### Le ciblage des données marché sur les polymères à cristaux liquides

(Début 1995)

#### Il y a 18 mots explicatifs pour le facteur 1 liste des mots a coordonnee negative

SEUILS POUR LA CONTRIBUTION 2\*MOY = 6.76 3\*MOY = 10.14

LIGNE	FACTEUR	QUALITE	CONTRI		
40	-0.67	99.09	14.72	3	VECTRA
41	-0.77	68.32	9.62	19	GIVEN
42	-0.87	92.47	13.61	14	APPLICATIONS
42	-0.88	47.02	7.37	46	RESISTANCE
43	-0.97	65.14	10.57	36	XYDAR
44	-1.04	170.41	21.46	11	CELANESE
44	-1.02	164.54	20.49	12	DETAILS
44	-0.97	69.96	11.80	32	PLASTICS
45	-1.06	127.81	15.79	24	HOECHST
45	-1.11	117.30	15.61	30	BRIEF
45	-1.10	74.77	10.84	47	COMPANY
45	-1.06	49.87	8.62	53	GRADES
45	-1.05	56.14	8.55	54	PRODUCTS
45	-1.09	52.39	7.53	68	MARKET
46	-1.15	48.56	7.49	75	DARTCO
47	-1.24	65.18	9.44	71	NOTED
47	-1.22	59.27	8.06	79	PLANT
48	-1.31	48.03	6.98	102	JAPAN

QUALITE = 67 ( 9 ) documents sélectionnés

## RESULTATS N°2 (suite)

### Exemple 1 : le ciblage des données marché sur les polymères à cristaux liquides

#### Liste des 9 documents

[Titres (champ "TI") et un exemple de résumé (champ "AB"); le champ SO signifie "source du document"] :

Dans *Résultats détaillés, Résultats n°2 p.176* : liste avec textes et surlignage des mots de la métaclé dans le texte.

**FACTEUR 1 Coordonnées Négatives au niveau de qualité : [ 67 , 250 [**

**Résumé** Document 00001044 ==> Ligne 41 Fac = -0.78 Qualité = 73.73 Contribution = 2.47

**TI : VECTRA LCP**

SO : Eur.Plast.News; 13, No.6, June 1986, p.46

**Résumé** Document 00000924 ==> Ligne 43 Fac = -0.95 Qualité = 69.40 Contribution = 1.82

**TI : VECTRA LCP MEETS CD REQUIREMENTS**

SO : Plast.Rubb.Wkly.; No.1285, 13th May 1989, p.12

**Résumé** Document 00000532 ==> Ligne 44 Fac = -1.02 Qualité = 97.77 Contribution = 7.75

**TI : LIQUID CRYSTAL POLYMER**

SO : Mod.Plast.Int.; 25, No.1, Jan.1995, p.48/50

AB : **Hoechst Celanese** has reaffirmed its commitment to the liquid crystal polymer segment by unveiling new capacity for its **Vectra** LCP. To meet the demand from fast-growing Asian markets, a 3,000 t/y **plant** will be built in **Japan** by Polyplastics, a Hoechst joint venture with Daicel Chemical Industries. Asian demand is expected to grow to approximately 910 tonnes by the end of the decade. DuPont has started construction of a commercial-scale facility at an existing site in Chattanooga and has named its melt-processable products Zenite. Amoco Polymers is targeting electronics applications with its newest grade, Xydar G-930. The glass-reinforced injection moulding resin is said to fill thin walls over long flow lengths with little or no flash. Despite their cost, LCPs are enjoying 25% annual growth worldwide and are forecasted to maintain that level through the decade. The global **market** for LCPs is about 4,500 tonnes.

**Résumé** Document 00001012 ==> Ligne 44 Fac = -1.02 Qualité = 80.65 Contribution = 2.33

**TI : LIQUID CRYSTAL POLYMER RATED FOR ELECTRONICS**

SO : Plast.S.Africa; 16, No.6, April 1987, p.49

**Résumé** Document 00001052 ==> Ligne 44 Fac = -0.99 Qualité = 71.37 Contribution = 4.25

**TI : MORE EUROPE INVOLVEMENT PROMISED BY CELANESE**

SO : Plast.Rubb.Wkly.; No.1134, 19th April 1986, p.9

**Résumé** Document 00001061 ==> Ligne 45 Fac = -1.11 Qualité = 71.60 Contribution = 2.76

**TI : LIQUID-CRYSTAL POLYMERS: NOW THEY ARE MELT-PROCESSABLE**

SO : Mod.Plast.Int.; 15, No.9, Sept.1985, p.66-8

**Résumé** Document 00000962 ==> Ligne 46 Fac = -1.20 Qualité = 67.56 Contribution = 2.95

**TI : NEW LIQUID CRYSTAL RESINS FOR E/E USES (E/E : electric/electronic)**

SO : Plast.Technol.; 34, No.10, Sept.1988, p.87/9

**Résumé** Document 00000964 ==> Ligne 47 Fac = -1.27 Qualité = 88.72 Contribution = 4.73

**TI : LIQUID-CRYSTAL POLYMER DEVELOPERS OFFER SYSTEMS WITH EASIER PROCESSING CHARACTERISTICS**

SO : Adv.Mat.(Newsletter); 10, No.10, 27th June 1988, p.2-3

[Résumé](#) Document 00000969 ==> Ligne 47 Fac = -1.22 Qualité = 72.33 Contribution = 3.01  
**TI : VECTRA LIQUID-CRYSTAL POLYMER TO BE PRODUCED HERE**  
SO : Jap.Chem.Week; 29, No.1474, 14th July 1988, p.3

Les documents rassemblés sont tous des documents commerciaux.

L'analyse de ces documents, le recoupement des informations et un classement par année montre que l'on s'engage vers :

- une augmentation de la production,
- un déplacement de la production vers l'Asie,
- une orientation vers des applications essentiellement électriques et électroniques.

L'article n°00000532 résume assez bien ce mouvement confirmé par les autres documents dès lors qu'on les envisage ensemble et chronologiquement.

Tel est l'exemple de ce que peut apporter une typologie thématique : non seulement le repérage de thèmes mais, dans le même temps, un ciblage rapide d'informations de synthèse sur un thème

On peut rétorquer que cette interprétation ne se fait que sur une partie de la base (l'association obtenue par analyse factorielle représentant finalement une sous-base). Nous tenterons d'évaluer en III-2 3 (p. 153) quelle est la perte d'informations. Cette interprétation sert de piste au même titre qu'une enquête et permet ensuite de refaire une recherche d'information de vérification sur l'ensemble des documents. En attendant elle dévoile une idée. Tel est le sens de la métaclé telle qu'elle révèle la cohérence des associations mises en évidence.

Un autre exemple de métaclé issu d'un traitement de documents en français (brèves issues du journal économique régional "Bref Rhône-Alpes") fait apparaître au facteur 1 positif un ensemble de documents concernant le thème de la formation. Tout comme le thème "marché" il s'agit d'un concept abstrait révélé par l'association *polytechnique, université, école, institut, formation ...*

Or la lecture du thème "formation" ainsi isolé met en évidence de façon frappante l'implication de l'INPG (Institut National Polytechnique de Grenoble) dans la quasi-totalité des articles, comme si, d'après les sources ici exploitées par "Bref Rhône-Alpes", l'INPG était impliqué dans la majeure partie des formations techniques de la région.

## RESULTATS N°3

### Exemple 2 : mise en évidence du thème "formation" dans l'analyse des brèves d'un journal économique régional "Bref Rhône-Alpes"

Il y a 36 mots explicatifs :

#### Pour le facteur 1 liste des mots a coordonnee positive

SEUILS POUR LA CONTRIBUTION 2\*MOY = 3.29 3\*MOY = 4.94

LIGNE	FACTEUR	QUALITE	CONTRI		
1	2.88	38.02	21.67	290	INV
1	2.89	235.43	23.52	446	POLYTECHNIQUE
1	2.89	175.92	33.97	586	UNIVERSITE
2	2.71	185.84	38.89	167	ECOLE
2	2.73	131.88	20.33	282	INGENIEURS
3	2.61	271.84	46.63	288	INSTITUT
7	2.01	57.84	8.89	305	JOSEPH
7	2.03	163.76	15.59	371	NATIONAL
9	1.67	104.32	7.72	201	ETAT
9	1.69	80.78	10.12	240	FORMATION
9	1.76	183.81	36.55	259	GRENOBLE
9	1.75	95.99	9.27	498	REGIONAL
9	1.66	63.65	8.39	601	VILLE
10	1.54	130.49	14.79	116	CONSEIL
10	1.62	28.49	6.04	444	POLYGONE
10	1.61	32.69	10.64	591	VALENCE
11	1.48	39.57	5.71	73	BUDGET
11	1.48	87.83	7.81	249	GENERAL
12	1.33	28.69	5.02	157	DIRECTEUR
13	1.11	129.47	16.24	23	ALPES
13	1.11	35.41	3.74	298	ISERE
13	1.10	125.69	11.51	497	REGION
14	1.05	17.59	3.30	109	COMMUNICATION
14	0.97	72.56	8.36	495	RECHERCHE
15	0.92	51.35	6.01	86	CENTRE
15	0.89	68.18	7.23	129	CREATION
15	0.83	55.15	6.61	186	ENTREPRISES
15	0.81	37.22	5.95	204	ETIENNE
15	0.89	116.52	17.01	326	LYON
16	0.73	36.13	6.33	156	DIR
16	0.73	42.49	4.28	481	PROJET
16	0.75	88.68	12.60	513	RHONE
17	0.49	54.86	3.92	206	ETRE
17	0.55	53.26	4.96	486	QUE
17	0.51	57.58	4.01	540	SONT
18	0.48	66.07	4.83	83	CE

QUALITE = 67' ( 48 ) doc.

## RESULTATS n°3 (suite)

Les premiers documents, parmi les 48 documents représentés sur le facteur 1 positif de l'analyse "Bref-Rhône-Alpes"

[Résumé](#) Document 00000997 ==> Ligne 4 Fac = 2.41 Qualité = 132.32 Contribution = 7.76

TITRE 00997 à INPG-UNIVERSITE JOSEPH FOURIER (GRENOBLE) :

DATE 30 Janvier 1991

CREATION D'HERCULE. L'INPG (Institut National Polytechnique) et l'Université Joseph Fourier de Grenoble créent Hercule, première formation doctorale européenne consacrée à l'étude de la matière condensée au moyen de grands instruments. Cette formation bénéficiera de la présence à Grenoble du Synchrotron de l'Institut Laue Langevin. Elle recevra pendant deux mois, 64 participants venant de 14 pays d'Europe sélectionnés parmi une centaine de candidats et 51 conférenciers européens.

[Résumé](#) Document 00001091 ==> Ligne 4 Fac = 2.39 Qualité = 152.49 Contribution = 8.23

TITRE 001091 GRENOBLE (38) : 22 MF POUR L'OBSERVATOIRE DES SCIENCES

DATE 23 septembre 1992;

22 MF POUR L'OBSERVATOIRE DES SCIENCES DE L'UNIVERS. L'observatoire des sciences de Grenoble vient de s'installer dans 2 300 m<sup>2</sup>, construits sur le domaine universitaire moyennant un investissement de 11,2 MF (CNRS : 5,9 MF ; Conseil régional : 2,7 MF ; Conseil général de l'Isère : 2 MF ; ministère de l'éducation nationale : 0,6 MF). Créé en 1985 par l'Université Joseph-Fournier (Grenoble I), l'Institut National Polytechnique de Grenoble (INPG) et l'Institut National des Sciences de l'Univers (INSU) du CNRS, cet observatoire est passé de 18 à 41 personnes en deux ans, bénéficiant de la politique de délocalisation du ministère de la recherche.

[Résumé](#) Document 00001200 ==> Ligne 4 Fac = 2.42 Qualité = 133.81 Contribution = 11.45

TITRE 001200 26 - L INPG ouvre

DATE 4 octobre 1995

L'INPG ouvre sa 9<sup>ème</sup> école d'ingénieurs à Valence. L'Institut national polytechnique de Grenoble (38) a ouvert, en partenariat avec la CCI de la Drôme, sa 9<sup>ème</sup> école d'ingénieurs à Valence Baptisée Esiar (Ecole Supérieure d'Ingénieurs en Systèmes Industriels Avancés Rhône-Alpes), elle prend le relais de l'Institut Supérieur d'Automatisme et de Robotique, créé à Valence par la CCI. L'Esiair va former des ingénieurs généralistes du génie électrique Les 67 premiers étudiants viennent d'entrer dans des locaux provisoires. Les locaux définitifs seront aménagés en plusieurs tranches Pour l'instant, 66 MF de travaux sont programmés. A terme, l'école pourrait accueillir jusqu'à 430 étudiants.

[Résumé](#) Document 00001136 ==> Ligne 5 Fac = 2.26 Qualité = 114.98 Contribution = 5.26

TITRE 001136 Cura (Rhône-Alpes) : Bernard Etlicher, nouveau pdt ...

[Résumé](#) Document 00000965 ==> Ligne 6 Fac = 2.18 Qualité = 150.48 Contribution = 7.34

TITRE 00965 38 - INPG : 140 MF prévus pour l'ENSPG et l'ENSERG ...

## RESULTATS n°3 (fin)

[Résumé](#) Document 00001087 ==> Ligne 6 Fac = 2.18 Qualité = 143.00 Contribution = 8.80

TITRE 001087 SAINT-ETIENNE (42) : un institut polytechnique régional ...

[Résumé](#) Document 00001088 ==> Ligne 6 Fac = 2.20 Qualité = 163.65 Contribution = 13.42

TITRE 001088 valence (26) : projet d'une école d'ingénieurs par ...

[Résumé](#) Document 00001129 ==> Ligne 6 Fac = 2.09 Qualité = 115.36 Contribution = 6.29

TITRE 001129 INPG/CCIIVD. : signature de la convention pour l'Esiair a ...

### III-2-1.2 Des métaclés à la typologie thématique de la base

Voici un exemple de typologie thématique pour "*Les polymères à cristaux liquides (LCP)*", à partir d'une base de 548 documents.

A chaque classe est associée en général une dizaine de documents, quelle que soit la façon de réaliser l'analyse factorielle. Nous donnons ci-après les principaux thèmes (le détail est donné dans *Résultats Détaillés : Résultats n°4* p. 179-192) :

1. Les données commerciales : voir précédemment Résultats n°2 (p.130-132).
2. Les données matériau : études des propriétés du matériau en fonction de la phase dans lequel il se trouve (smectique, nématique , cholestérique), ces phases se caractérisent par l'arrangement des molécules entre elles. Les polymères à cristaux liquides sont en effet des liquides qui s'arrangent comme des solides : ils ont une structure ordonnée. Leurs propriétés dépendent de cette structure.
3. Propriétés, dans la transformation, des mélanges ("blends") à base de polymères à cristaux liquides. Les LCP sont des matériaux chers et l'idéal est donc de pouvoir les mélanger avec des matériaux moins onéreux tout en obtenant des propriétés qui demeurent intéressantes.
4. LCP et fibres en particulier fibres de carbone.
5. Le problème des lignes de soudure dans la fabrication des pièces en LCP.
6. Les stratifiés ("laminates")
7. Les LCP utilisés comme films barrière
8. Organisation en lamelles des films extrudés.

### III-2-1.3 Les déceptions de l'interprétation :

Nous avons rencontré essentiellement trois motifs de déception dans les résultats obtenus :

1. Ce que nous appelons "les associations triviales" : l'AFC ne semble pas dans ce cas apporter d'élément d'interprétation.
2. Une correspondance non essentielle entre les mots de la métaclé et les documents de la même métaclé, les mots justifiant de la cohésion de l'association des documents n'apparaissant pas.
3. La difficulté et la fécondité de la notion "d'hyperplan" d'un document : comment un document apparaissant sur plusieurs facteurs (c'est-à-dire ayant plusieurs dimensions) peut

tantôt renvoyer à un problème de représentation tantôt illustrer la richesse d'interprétation possible du document.

Voici plus précisément en quel sens ces difficultés, mais aussi certains aspects positifs qui leur sont attachés, nous sont apparus. Nous reprenons ci-après les trois points précédemment évoqués.

#### 1- "les associations triviales"

Dans le cas de bases hétérogènes, c'est-à-dire de base dont les documents ont été collectés de façon disparate, avec des mots-clés différents, il n'est pas surprenant d'obtenir des associations "non conceptuelles", apparemment triviales. Dans l'exemple ci-dessous concernant le journal "Bref Rhône-Alpes", nous trouvons tous les documents concernant l'entreprise Caty Polymères regroupés sur deux facteurs. Aucun croisement avec des documents ne concernant pas cette entreprise ne s'est effectué. On a retrouvé ce phénomène à plusieurs reprises dans l'analyse de cette base.

Toutefois on peut se poser la question de savoir pourquoi un tel regroupement s'est effectué. Si cette association a été constituée en métaclé c'est probablement parce qu'elle révèle une singularité par rapport à l'ensemble de la base. Il reste que la possibilité de séparer des éléments de nature différente par facteur n'apporte pas d'informations supplémentaires. Ce phénomène traduit une fragmentation de la base en sous-ensembles singuliers n'ayant pas de point communs. Nous présentons ce résultat ci-après.

### **RESULTATS N°5**

#### **Exemple d'associations triviales**

Journal "Bref Rhône-Alpes"

#### **FACTEUR 1 Coordonnées Négatives au niveau de qualité : [ 67 , 250 [**

[Résumé](#) Document **00000180** ==> Ligne 29 Fac = -1.19 Qualité = 71.66 Contribution = 4.01

**TITRE 00180 Caty polymères (Bourg-de-Péage) : 1,2 MF de**  
**NUMERO 1262 . DATE 19 janvier 1994;**

TEXTE Caty polymères (Bourg-de-Péage) : 1,2 MF de ; résultats Caty Polymères, fabricant de pièces en caoutchouc et polyuréthanes, a enregistré, en 93, un CA de 57, 5 MF (+ 18,5 %), dont 13 % hors France, un RN de 1,2 MF (237 000 F en 92) et un cash flow de 4 MF. Elle a investi 4 MF et consacré la même somme à la recherche appliquée. Par rapport à 88, date du rachat de l'entreprise par Fernand Dolisy (actuel pdg), le CA a progressé de 58 %. L'entreprise emploie 115 personnes dans son usine de 5 000 m<sup>2</sup> à Bourg-de-Péage et envisage de construire une seconde unité de 1 200 m<sup>2</sup> (investissement : 1,2 MF) mitoyenne à l'actuelle. Caty Polymères est la nouvelle raison sociale de l'entreprise issue de la récente fusion entre les sociétés Caty SA et Société nouvelle Caty SA. Cette opération a permis une simplification des structures et un renforcement du niveau des fonds propres avec un capital de 20 MF.

**Résumé Document 0000233 ==> Ligne 30 Fac = -1.36 Qualité = 87.03 Contribution = 2.65**  
**TITRE 00233 Caty Polymères (26) : zénith de la réussite de la Drôme**  
**NUMERO 1303. DATE 14 décembre 1994;**

TEXTE Caty Polymères (26) : zénith de la réussite de la Drôme. La société de Bourg-de-Péage, Caty Polymères, 111 salariés (pdg : Fernand Dolisy), vient de recevoir le Zénith d Or de la réussite du département de la Drôme. Spécialisée dans la fabrication de pièces en caoutchouc et polyuréthane, elle est passée entre 1988 et 1993 de 37 MF de CA à 59 MF. Sa part du CA à l'export s'est établie, en 93, à 20 %, son résultat net à 1,9 MF La R & D représentait 7 % du CA. Elle a consacré 15,6 MF aux investissements de 1989 à 1993.

**Résumé Document 0000254 ==> Ligne 30 Fac = -1.35 Qualité = 91.82 Contribution = 5.55**  
**TITRE 00254 26**  
**NUMERO 1347. DATE 29 novembre 1995;**

TEXTE ; Caty Polymères dans sa nouvelle usine. La société Caty Polymères, fabricant de pièces en caoutchouc et en polyuréthane, vient d inaugurer, sur la ZI de Bourg-de-Péage, une unité de 1 300 m2, dédiée au polyuréthane. Elle représente un investissement de 3,7 MF (Bref 1317). En 1995, la société a également investi 2,65 MF dans l'achat de matériels. Caty Polymères (pdg : Fernand Dolisy) a enregistré, en 94, un CA de 62 MF avec 110 salariés. Caty Polymères est coiffée par FD Holding, qui contrôle également Caty 3M, spécialisée dans les modèles de fonderie, maquettes et moules (12 salariés) à Bourg-de-Péage. Le groupe, présidé par Fernand Dolisy, a enregistré au 30/10/95 une progression de 12 % de son CA cumulé par rapport à la même période de 94. Les nouveaux produits à forte valeur ajoutée ont augmenté de 43 %, représentant désormais près de la moitié du CA. L'entreprise s'est dotée d'un Comité de Direction.

**Résumé Document 0000293 ==> Ligne 37 Fac = -2.29 Qualité = 108.70 Contribution = 24.95**  
**TITRE 00293 26/87**  
**NUMERO 1370 DATE 29 mai 1996;**

TEXTE ; Caty Polymères s'adosse au groupe Alma. Le groupe limousin AG-Finances-Groupe Alma vient de prendre le contrôle à 100 % de Caty Polymères, spécialiste de la formulation de caoutchouc haut de gamme et d élastomères de polyuréthane, basée à Bourg-de-Péage (26). Guy Maurette, pdg de AG Finances, et Alain Magne, dg, deviennent respectivement pdg et dg du groupe drômois rebaptisé Alma Caty. Quant à Fernand Dolisy (ex-pdg du groupe Caty Polymères), il rejoint AG Finances comme vice-président. En 8 ans le petit groupe FD Caty chapeautant les 2 sociétés, Caty Polymères et Caty (3 MF, spécialisée dans les modèles de fonderie, maquettes et moules), est passé de 37 MF de CA à près de 61 MF en 95 (dont 15 % à l export) avec 130 personnes En 8 ans, 25 MF d investissements ont été réalisés dans l'outil de production regroupé dans plusieurs bâtiments sur la ZI de Bourg-de-Péage (Bref 1347) Le groupe Alma-AG-Finance, domicilié à Isle-sur-Vienne (87), et spécialisé dans la fabrication de pièces techniques thermoplastiques pour l'industrie, est majoritairement contrôlé par ses fondateurs (Guy Maurette et Alain Magne). Il bénéficie de la présence à son capital de l'équipementier automobile MGI-Coutier (Champfromier) Le groupe Alma-AG-Finances chapeaute 4 filiales : Alma Industrie (Aixe-sur-Vienne), Alma-Est (Dampierre-les-Bois), Alma-Sermi (Bellignat) et Alma-Saint-Gérons (Lorp-Sentaraille) Il a enregistré en 1995 un CA/HT de 65 MF (52,6 MF en 94) avec 126 salariés. En 1996, son objectif est d atteindre 90 MF (sans Alma Caty) et de dépasser les 150 MF avec Alma Caty. Alma Caty vise 62 MF de CA en 96, 67 MF en 97 et 72 MF en 98.

## **JOURNAL**

### **FACTEUR 2 Coordonnées Positives au niveau de qualité : [ 250 , 500 ]**

**Résumé Document 00001098 ==> Ligne 31 Fac = 4.05 Qualité = 296.89 Contribution = 15.32**  
**TITRE 001098 Caty Polymères (Bourg-de-Péage) : le catane élargit**  
**NUMERO 1210. DATE 25 novembre 1992;**

TEXTE Caty Polymères (Bourg-de-Péage) : le catane élargit ses applications. Caty Polymères a élargi les applications de son catane (matériau viscoélastique pour l'amortissement anti-vibratoire mis au point en partenariat avec Rhône-Poulenc et l'Institut biomédical des sports de vitesse) après les ...

**Résumé Document 00000006 ==> Ligne 34 Fac = 3.31 Qualité = 336.56 Contribution = 21.01**

**TITRE 006 Caty polymères (Bourg-de-Péage)**  
**NUMERO 1317.** DATE 29 mars 1995;

TEXTE Caty polymères (Bourg-de-Péage); investit dans 1 300 m2 supplémentaires. Caty Polymères (pdg : Fernand Dolisy), fabricant de pièces en caoutchouc et polyuréthane, vient d'engager les travaux de construction de 1 300 m2 de bâtiments s ajoutant aux 5 000 m2 actuels Cet investissement de 3 MF va lui permettre d'accueillir ses activités polyuréthane. Une trentaine d emplois devraient être créés à terme. Caty Polymères, 111 salariés, a réalisé en 94 un CA de 61,5 MF (59 MF en 93).

2- Une correspondance non essentielle entre les mots de la métaclé et les documents de la même métaclé

La seconde déception que nous avons rencontrée est liée à ce que la sélection des mots en fonction d'un seuil de contribution induit parfois qu'ils ne reflètent pas le contenu des documents sélectionnés. Ils ne peuvent pas jouer alors leur rôle de "guide de lecture" des textes associés car le seuil de sélection choisi n'a pas permis de faire apparaître ce qui a rassemblé les documents.

Par exemple dans le facteur 10 positif de l'une des analyses *Polymères à cristaux Liquides* (LCP 296) , les mots *temperature*, *temperatures* et *DSC (Differential Scanning Calorimetry)* n'apparaissent pas dans la métaclé. ***Ils sont pourtant le point commun de tous les documents de cette métaclé.*** Seul le mot de *Treatment* est significatif à la fois dans la métaclé (il a la plus forte contribution à la valeur propre) et dans les documents.

Ce phénomène s'explique par deux raisons :

- le mot n'est pas sélectionné parmi ceux avec lesquels l'analyse s'est effectuée. Nous reviendrons ci-après III-2. 2 (p. 145-153) sur les manières de sélectionner les mots.
- Sa contribution à la valeur propre est inférieure au seuil fixé.

## RESULTATS N°6

Exemple de difficulté d'interprétation à partir des mots de la métaclé :

**L'absence des mots *température* et *DSC (Differential Scanning Calorimetry)***

**Il y a 25 mots explicatifs :**

### LISTE DES MOTS A COORDONNEE POSITIVE

SEUILS POUR LA CONTRIBUTION 2\*MOY = 6.76 3\*MOY = 10.14

LIGNE	FACTEUR	QUALITE	CONTRI		
1	4.18	66.78	43.74	295	TREATMENT
9	2.97	80.60	29.40	214	PEAK
9	2.91	41.16	28.21	222	TPB
10	2.78	69.56	25.72	215	PREPREGS
13	2.40	68.42	28.96	148	LAMINATES
13	2.39	36.67	16.63	235	ISOTHERMAL

15	1.98	51.98	13.03	205	DOUBLE
15	2.02	17.24	11.93	225	ALKENE
16	1.94	35.16	12.50	207	FLASH
16	1.92	41.96	12.30	224	UNIDIRECTIONAL
16	1.84	25.37	8.44	278	MOIETIES
17	1.65	92.60	36.39	60	PPS
17	1.72	71.36	22.23	98	CRYSTALLISATION
17	1.79	27.89	9.36	236	KEVLAR
17	1.67	32.47	8.16	241	OXY
18	1.61	86.29	26.91	77	MELTING
19	1.42	34.46	8.47	186	VENTURE
19	1.43	27.17	7.71	191	HYDROXY
20	1.23	43.64	8.82	131	JOINT
20	1.22	24.78	7.49	154	TONNES
23	0.88	19.68	7.04	84	STRESS
24	0.71	51.14	21.16	8	FIBRES
24	0.63	30.59	7.57	39	HEAT
24	0.69	19.06	6.88	52	TRANSITION
26	0.38	26.68	8.66	4	FIBRE

QUALITE = 67 ( 5 ) doc

### RESULTATS N°6 (suite) : Les 5 documents :

Résumé N0 00000825

#### TI - ISOTHERMAL HEAT TREATMENT OF A THERMOTROPIC LCP FIBRE

AU - Sarlin J; Tormala P ;

SO - J.Polym.Sci.Polym.Phys.; 29, No.4, 30th March 1991, p.395-405

AB - The effect of heat **treatment** on as-spun fibres made of a commercial liquid crystalline polymer (Vectra A900) was studied by tensile measurements and **DSC**. The heat treatment increased fibre strength considerably and also elongation at break but had a weaker effect on modulus. **DSC** showed melting behaviour to be strongly dependent on **treatment** temp., the melting curve displaying two endotherms whose **temperature** increased with **treatment** time.

Résumé N0 00000848

#### TI - MELTING BEHAVIOUR OF A THERMOTROPIC LCP (VECTRA A900) FIBRE

AU - Sarlin J; Tormala P ;

SO - Polym.Comm.; 31, No.10, Oct.1990, p.372-4

AB - The melting behaviour of thermotropic liquid crystalline polymer as-spun fibres was investigated by **DSC**. The fibres displayed a double melting peak with an exotherm occurring close to the lower **temperature** peak. This exotherm and double melting peak were not discernible in the second heating run, indicating that they were closely connected to the thermal and stress history of the sample. 20 .

Résumé N0 00000608

#### TI - SELF-REINFORCED LCP-LCP PREPREGS AND LAMINATES

AU - Ding R; Isayev A I ;

CP - Antec '93. Conference Proceedings

SO - New Orleans, La., 9th-13th May 1993, Vol.II, p.1176-82. 012

AB - Self-reinforced prepregs were prepared by stretching sheets made of blends of two liquid crystal copolyesters, i.e. Vectra A950 (Hoechst Celanese) and Ultrax KR 4003 (BASF), exiting from a coathanger die attached to a static mixer/single-screw extruder set-up. Processing at this stage was carried out at **temperatures** at which both components were melt processable. The prepregs were laid up in multi-layers and compression moulded into quasi-isotropic or unidirectional laminates at **temperatures** below the melt processing **temperature** of one component but above the melting point of the other. The mechanical properties of the prepregs and laminates were measured. The properties of quasi-isotropic laminates were

calculated based on those of unidirectional laminates and compared with measured properties. This led to the conclusion that some damage occurred during consolidation of quasi-isotropic laminates. 14 .

RESUME NO 00000504

**TI - PRELIMINARY INVESTIGATION OF FLASH FORMATION DURING INJECTION MOULDING OF POLYPHENYLENE SULPHIDE(PPS) AND LIQUID CRYSTALLINE POLYMER(LCP) BLENDS**  
AU - Shonaik G O; Yamaguchi S; Ohta M; Hamada H; Maekawa Z; ; Nakamichi M; Kosaka W; Toi K  
SO - Polym.Engng.Sci.; 35, No.3, Mid-Feb.1995, p.240-6

AB - Flashing in pure PPS and a blend containing PPS and LCP during injection moulding was studied by **DSC** and SEM. The shape of the flash was observed by use of a projector. Flashing was detected in pure PPS and 90/10 PPS/LCP blend but was not found in other compositions, including pure LCP. The **DSC** thermograms of the flash revealed both exothermic and endothermic peaks at around 120 and 285C. The first peak, known as **crystallisation temp.** on heating, occurred as a result of early crystallisation of PPS. The observed double peaks indicated that the degree of crystallinity was lower in the flash than in the moulded part. The morphological studies revealed the presence of LCP fibrils in the skin region and droplets in the core region of 90/10 PPS blend. The absence of flash was attributed to the diameters of the fibrils and droplets, which were found to increase with increasing LCP component.

### RESULTATS N°6 (fin)

Résumé NO 00000797

**TI - CRYSTALLISATION BEHAVIOUR OF POLYPHENYLENE SULPHIDE**

AU - Budgell D R; Day M ;

SO - Polym.Engng.Sci.; 31, No.17, Mid-Sept.1991, p.1271-8

AB - The crystallisation behaviour of polyphenylene sulphide (PPS) has been examined by **DSC** as a function of melt **temperature**, residence time in the melt, and the presence of a liquid crystalline polymer. Thermal history was found to play a significant part in the determination of crystalline kinetics. Short residence times at low melt **temperatures** and high melt **temperatures** alone resulted in a low value for the Avrami exponent. The former effect was ascribed to incomplete melting of the polymer while the latter effect was attributed to thermal degradation. Blending a liquid crystal polymer, Vectra A950, with PPS had only a minor effect on the crystallisation behaviour. 46 .

3- Difficulté et fécondité de la notion d'hyperplan d'un document.

La multidimensionnalité de l'analyse factorielle des correspondances et l'unidimensionnalité de la représentation par métaclé que nous venons d'exposer peuvent conduire à des difficultés d'interprétation bien connues dans l'usage de ce type de méthodes. Ainsi obtient-on parfois des regroupements de mots et de documents n'ayant rien de commun sur une métaclé alors qu'une représentation à deux dimensions élucide cette incohérence en séparant les groupes de données.

Toutefois les mêmes raisons qui expliquent la difficulté de représentation liée à l'aspect multidimensionnel de l'analyse factorielle sont aussi celles qui sont susceptibles de rendre l'interprétation féconde. Ainsi peut-on non seulement interpréter des associations de mots et de documents (des thèmes) mais aussi ce que l'on peut appeler l'*hyperplan d'un document* c'est-à-dire l'ensemble des métaclés (et donc des thèmes) sur lesquelles il est représenté. Un document peut être monothématique s'il n'est présent que sur une dimension, mais peut être multithématique : son interprétation peut alors se faire soit par rapport à sa spécificité ou par rapport à l'ensemble

des thèmes qui le constituent. Il s'envisage alors à partir de contextes différents constitués de mots et de documents associés différents.

Il en est ainsi pour le document 99, que nous citons ci-dessous (Résultats n°7). Ce document fait partie de l'une de nos analyses sur les capteurs (analyse n°3 ; cf p.210). Il est représenté à la fois sur le facteur 1 et le facteur 4. Il est envisagé sous deux aspects différents : sur le facteur 1 il renvoie au contexte des capteurs impliqués dans la transformation des composites; sur le facteur 4 il fait partie du thème général relatif au contrôle des process (du point de vue des matières /composites, caoutchouc, thermoplastique.../ ou des process /injection, extrusion).

<b>RESULTATS N°7</b>
<p>HYPERPLAN D'UN DOCUMENT :</p> <p>Le document 99 dans l'étude <i>Capteurs dans les procédés de transformation des polymères</i>            Sa double appartenance au facteur 1 négatif et au facteur 4 négatif</p>
Résumé NO 00000099
<b>@TI IN-SITU SENSOR MONITORING AND INTELLIGENT CONTROL OF THE RESIN TRANSFER MOULDING PROCESS</b>
<b>@AU Kranbuehl D E; Kinsgley P; Hart S; Hasko G; Dexter B; Loos A C;</b>
<b>@SO Polym.Composites; 15, No.4, Aug.1944, p.299-305</b>
<b>@DT Journal</b>
<b>@LA English</b>
<p>@AB An intelligent closed-loop expert control system was developed for automated control of the resin transfer moulding process of a graphite fibre preform using an epoxy resin, E905L. The sensor model system was developed to make intelligent decisions based on the achievement of landmarks in the cure process, such as full preform impregnation, the viscosity, and the degree of cure of the resin rather than time or temperature. In-situ frequency dependent electromagnetic sensor and the Loos resin transfer model are used to monitor and control the processing properties of the epoxy resin during resin transfer moulding impregnation and cure of an advanced fibre architecture stitched preform. 13</p>
<p><b>@DE AUTOMATION ; CARBON FIBRE- REINFORCED PLASTIC ; CFRP ; CLOSED LOOP ; COMPANY ; COMPOSITE ; CURING ; DATA ; DEGREE OF CURE ; EPOXIDE RESIN ; EPOXY RESIN ; EQUATION ; FLOW CHART ; FREQUENCY DEPENDENT ELECTROMAGNETIC SENSOR ; GRAPH ; GRAPHITE FIBRE-REINFORCED PLASTIC ; IMPREGNATION ; IN-SITU ; INSTITUTION ; INTELLIGENT PROCESSING ; MODEL ; MONITORING ; PICTURE FRAME ; PLASTIC ; PREFORM ; PROCESS CONTROL ; PROPERTIES ; REINFORCED PLASTIC ; RESIN TRANSFER MOULDING ; SENSOR ; TABLES ; TECHNICAL ; TEMPERATURE ; THEORY ; THERMOSET ; TIME ; VISCOSITY</b></p>
<b>@PY 1994</b>

Les même termes (communs aux deux facteurs) que sont *cure* et *resin* sont associés :

- dans le facteur 1 à des noms de matière de *fibres*, *epoxy* (thème : composite et thermodur) et à *model* et *dielectric* montrant que l'enjeu est ici la modélisation des propriétés de ce type de matière lors de la transformation (dans ce document 99, il s'agit du Resin Transfer Moulding process). C'est le thème de l'ensemble des documents de ce facteur, dont nous présentons les titres ci-après.

- dans le facteur 4 à ceux de *control*, *developments*, *automation* illustrant le thème du "process control".

### RESULTATS N°7 (suite)

HYPERPLAN DU DOCUMENT N°99 :

**Le document 99 s'exprime à la fois sur deux métaclés (facteur 1 positif, facteur 4 négatif) exprimant par là deux thèmes**

**IL Y A 8 MOTS EXPLICATIFS :**

**POUR LE FACTEUR 1 LISTE DES MOTS A COORDONNEE NEGATIVE**

LIGNE	FACTEUR	QUALITE	CONTRI		
35	-1.09	177.88	32.16	36	DURING
35	-1.10	115.79	25.93	43	FIBRE
36	-1.13	87.79	22.09	69	MODEL
43	-1.68	135.54	47.00	33	DIELECTRIC
43	-1.64	337.11	83.32	104	RESIN
47	-1.93	279.10	43.35	114	SITU
49	-2.06	323.88	64.44	40	EPOXY
50	-2.16	531.77	152.32	22	CURE
		<u>QUALITE = 250</u>	' ( 8 ) doc		
		<u>QUALITE = 67</u>	' ( 43 ) doc		

**IL Y A 5 MOTS EXPLICATIFS :**

**POUR LE FACTEUR 4 LISTE DES MOTS A COORDONNEE NEGATIVE**

LIGNE	FACTEUR	QUALITE	CONTRI		
37	-0.49	119.84	46.90	21	CONTROL
39	-0.74	62.63	29.49	22	CURE
39	-0.72	64.27	26.11	104	RESIN
45	-1.39	85.30	39.45	31	DEVELOPMENTS
46	-1.52	74.13	33.49	6	AUTOMATION
		<u>QUALITE = 67</u>	' ( 9 ) doc		

### III-2-1.4 Les cartographies thématiques

L'avantage des cartographies est d'abord de présenter sous forme synthétique plusieurs thèmes. La seule vision d'un plan à deux dimensions permet de voir quatre thèmes simultanément. La principale difficulté est de représenter tous les points (mots et documents) sur une carte. Les produits logiciels du marché présentent en général des cartes illisibles au premier regard mais qui comportent des fonctions "zoom" permettant d'agrandir le nuage de point et des fonctions hypertexte pour obtenir "le contenu du point". Nous avons, pour notre part, établi des cartes à partir des points sélectionnés dans les métaclés, uniquement pour les points-mots. Pour obtenir les documents il suffit de se reporter à la métaclé.

Dans l'exemple ci-après nous montrons une carte des deux premiers facteurs de l'analyse *Polymères à cristaux liquides* (analyse n°6-3). Cette représentation à deux dimensions permet une superposition des métaclés (facteur 1 positif et facteur 2 positif) ayant l'une et l'autre comme objet la structure de la matière (l'essentiel des documents du facteur 2 positif faisant d'ailleurs partie du facteur 1 positif). Ce thème occupe le cadran droit de la carte. L'avantage de cette représentation à deux dimensions est de mettre en évidence, dans ce quadrant droit, des sous-thèmes que nous avons entourés : les termes ayant trait à la structure du matériau *stricto sensu* (*chain, nematic, smectic...*), les termes décrivant une propriété physique afférente à la structure (*field, electric* : champ électrique), termes relatifs à la cinétique de polymérisation (*theory, kinetics, polymerisation, transition...*), ceux concernant l'ordre moléculaire (*molecular, order, phase*).

Ces associations de termes ne sont visibles que grâce aux deux dimensions. Elles forment des guides de lecture du même sous-ensemble de documents que ceux mis en évidence dans l'interprétation linéaire (par métaclé). Mais elles permettent une lecture plus fine et plus rapide (les deux métaclés se superposant visiblement) de la même analyse.

## RESULTATS N°8

FACTEUR 1 ET FACTEUR 2 DE L'ANALYSE POLYMERES A CRISTAUX LIQUIDES  
(LCP 296 MOTS)

LES INDIVIDUS VERIFIENT AU MOINS UNE DES CONDITIONS SUIVANTES :

c\_\_1 =[ 6.757 1000.000]

c\_\_2 =[ 6.757 1000.000]

**LAXE HORIZONTAL(f\_1)--AXE VERTICAL(f\_2)**

NOMBRE DE POINTS : 57

==ECHELLE : 4 CARACTERE(S) = 0.22 1 LIGNE = 0.09

-----TPB +		1
	<i>MATIERE: Etude de la</i>	2
	<i>structure des LCP</i>	3
		4
	SILOXANE METHACRYLATE	5
	MESOGENIC	6
		7
	<b>SIDE DIRECTOR</b>	8
		9
		10
	<b>CHAIN</b>	11
	GROUPS SMECTIC	12
	<b>NEMATIC</b>	13
		14
	FIELD	15
		16
	ELECTRIC	17
	THEORY KINETICS	17
	POLYMERISATION	18
		19
	TRANSITION	20
		21
JAPAN		21
	NOTED	22
	BRIEF	22
	MARKET DETAILS	23
	COMPANY	24
	CELANESE	25
	HOECHST APPLICATIONS GIVEN	26
	MOLECULAR	26
	ORDER	27
	PHASE	28
	VECTRA	29
-----+		30
	<i>MATERIAU :</i>	31
	<i>Etude des mélanges à base de LCP</i>	32
	<i>Propriétés rhéologiques</i>	33
		34
	INJECTION	35
		36
		37
		38
		39
	PPS LCP	40
	PRESSURE POLYCARBONATE VISCOSITY	41
	PBTP SHEAR	42
	<b>BLENDS</b> CRYSTALLISATION	43
	PEEK ELONGATIONNAL	44
	FIBRILS	45
	VISCOSITIES	46
	TLCP	47
		48
		49
		50
		51
		52
		53
		54
	PPO	55
-----+		56

NOMBRE DE POINTS SUPERPOSES : 11

Les mots ayant une forte contribution apparaissent en gras.

ALKENE(DIRECTOR) MAGNETIC(SMECTIC) SPACERS(GROUPS) CHAIN(FIELD)

PLANT(NOTED) PRODUCTS(DETAILS)DARTCO(MARKET) PLASTICS(CELANESE) XYDAR(CELANESE) RESISTANCE(APPLICATIONS)

GRADES(HOECHST)

L'avantage des cartographies ajouté à celui de la représentation à trois dimensions permises par les ordinateurs d'aujourd'hui sont susceptibles d'améliorer les problèmes de représentations liées à l'analyse factorielle des correspondances c'est-à-dire à en fait à la représentation multidimensionnelle. Plus une représentation comportera de dimensions moins l'analyse révélera d'ambiguïtés, mais plus elle deviendra difficile à comprendre et à regarder. En d'autres termes on retrouve la difficulté que la représentation géométrique tente de lever... La représentation linéaire est la seule permettant de lire l'ensemble des points mots et documents sélectionnés.

### III-2-2 INFLUENCE DU PRINCIPE DE SELECTION DES UNITES STATISTIQUES SUR LES RESULTATS

Le traitement de l'information textuelle repose sur :

- le découpage des textes en unités documentaires, unités au sein desquelles s'effectuent les comptages des unités linguistiques.
- la sélection des unités linguistiques prises en compte dans le traitement.

La question du découpage des textes vaut surtout pour le traitement du texte intégral. Il est en effet possible de compter une unité linguistique soit dans un document entier soit dans des unités logiques plus fines : le chapitre, le paragraphe, la phrase.

Pour notre part nous avons traité le résumé des articles car c'est la seule forme de document électronique disponible en matière d'information scientifique et technique. Le résumé est pris comme une seule unité documentaire. Le traitement du résumé a ce double avantage : celui de traiter une unité textuelle (au lieu de comptabiliser des descripteurs) et celui de laisser envisager ce que peut être le traitement des paragraphes dans les textes intégraux.

La sélection des unités linguistiques peut se faire sur deux critères :

- le niveau de complexité : de la forme non lemmatisée au syntagme nominal.
- la fréquence du mot ou de son occurrence et plus généralement un critère quantitatif de sélection. La fréquence est celle du mot dans l'ensemble de la base; celle de l'occurrence est le nombre de documents ayant ce mot.

En ce qui concerne la complexité on peut dire que plus on prend en compte une unité linguistique complexe, plus on gagne en précision et plus on perd de liens entre les unités. Paradoxalement la précision occasionne une perte d'informations.

La sélection des mots a pour but de rendre possible la comparaison entre les documents. La recherche d'information repose sur la notion de similarité (ou de distance) entre des énoncés : il s'agit d'apparier le texte d'une requête avec celui des documents ou de classer les documents en fonction de leur contenu. Pour que la comparaison soit possible il faut des éléments communs : si tous les documents étaient constitués d'hapax<sup>48</sup> ils ne seraient pas comparables. Seule la recherche d'appariement d'une chaîne de caractères avec une autre comme si l'on recherchait un nom propre ou l'introduction d'un lexique pour traduire en synonymes les termes existants avec ceux de la requête seraient envisageables.

Ici nous cherchons à détecter une typologie thématique. Pour cela, il faut à la fois éliminer les hapax et les mots très fréquents c'est-à-dire les mots singularisant chaque document ou ceux, au contraire, qui les indifférencient et constituent ceux par quoi ils sont tous semblables. Ce dernier cas est en général celui des déterminants (articles, pronoms essentiellement). La nature de ces termes "indifférenciateurs" peut changer selon la nature des textes traités.

De plus, dans les textes scientifiques, cette catégorie de termes est tout autant représentée par les mots décrivant le sujet. Par exemple, dans l'analyse d'une base de données sur les polymères à cristaux liquides, le mot "polymère à cristaux liquides" peut être mis sur le même plan que les déterminants, c'est-à-dire considéré comme un mot vide puisque tous les documents auront ce mot. Le problème est donc de déterminer le seuil à partir duquel vont être sélectionnés les mots que nous pouvons qualifier de non discriminants.

Dans l'expérience suivante nous avons comparé les résultats de plusieurs analyses factorielles issues de différentes façons de sélectionner les mots des documents de la base. L'objectif est de mesurer les différences entre les typologies thématiques obtenues en fonction des procédés de sélection des mots utilisés.

On a comparé, à partir *de la même base de documents* les AFC effectuées sur :

---

<sup>48</sup> Un hapax est un mot n'apparaissant qu'une seule fois.

-les mots ayant une fréquence supérieure ou égale à 1% du nombre des documents de la base. Cette analyse porte dans l'exemple qui suit sur 746 mots différents : **nous l'appellerons l'analyse 746.**

-les mots ayant une fréquence d'occurrence supérieure ou égale à 1% du nombre des documents de la base. La fréquence d'occurrence comptabilise le nombre de documents dans lequel le mot apparaît et non plus le nombre de fois qu'il apparaît dans la base. Cette analyse repose ici sur 673 mots : **nous l'appellerons l'analyse 673.**

- Les mots issus de la première analyse (AFC croisant les documents de la base avec les mots ayant une fréquence supérieure à 1% du nombre des documents) dont la contribution à la valeur propre est supérieure ou égale à 2\* la moyenne de l'inertie totale. Cette analyse a été ainsi effectuée sur 296 mots : **nous l'appellerons l'analyse 296.**

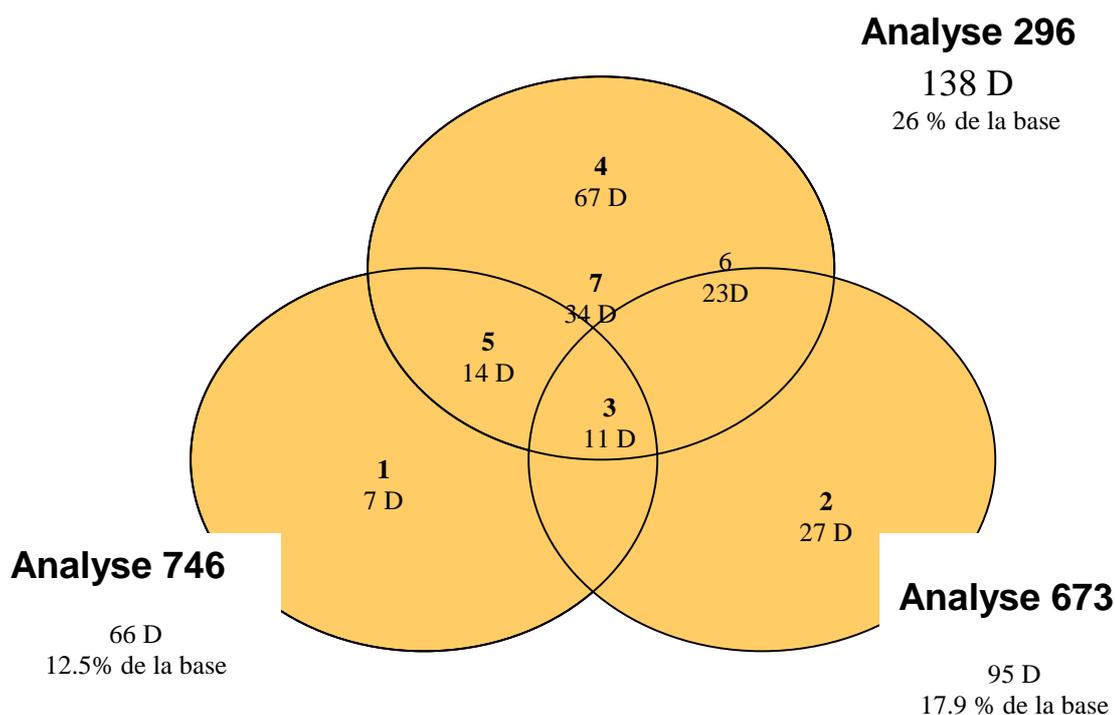
### **III-2-2.1 Comparaison des trois analyses entre elles :**

Le schéma "Comparaison des trois analyses entre elles" résume :

- le nombre de documents exprimés sur les trente premiers facteurs de chaque analyse
- le nombre de documents communs
- le nombre de documents différents.

## RESULTATS N°9

### COMPARAISON ENTRE LES TROIS ANALYSES POLYMERES A CRISTAUX LIQUIDES



*De 1 à 7 : numéros des zones*

*n D = nombre de documents*

ZONE 1 : documents propres à l'analyse 746 (7 documents)

ZONE 2 : documents propres à l'analyse 673 (27 documents)

ZONE 3 : documents communs à l'analyse 746 et à l'analyse 673 (11 documents)

ZONE 4 : documents propres à l'analyse 296 (67 documents)

ZONE 5 : documents communs à l'analyse 746 et à l'analyse 296 (14 documents)

ZONE 6 : documents communs à l'analyse 673 et à l'analyse 296 (23 documents)

ZONE 7 : documents communs aux trois analyses (34 documents)

### **III-2-2.2 Comparaison des typologies thématiques obtenues avec l'analyse 1 (746) et l'analyse 2 (673)**

Les typologies thématiques obtenues sont identiques jusqu'au facteur 3 positif inclus : seul diffère le nombre de documents sélectionnés. Les mêmes mots apparaissent et les plus fortes contributions du facteur concernent les mêmes mots. Les documents sélectionnés dans l'une des analyses forment un sous-ensemble de celui sélectionné dans l'autre.

Pour chacune des deux analyses, 23 documents sont communs sur 63 cités dans les dix premiers facteurs, soit 36% des documents communs.

Les différences entre les deux typologies commencent à partir du facteur 3 négatif où l'on ne retrouve plus la logique d'inclusion d'un sous-ensemble de documents dans l'autre quand on compare les deux analyses. En revanche les thèmes varient assez peu.

Voici quelques exemples du type de variation que l'on peut constater.

Dans le facteur 3 négatif de chaque analyse les deux mots communs "injection" et "moulding" font partie des quatre mots ayant la plus forte contribution à la valeur propre et la métaclé concerne bien, dans les deux cas le moulage par injection des LCP.

- Dans le cas de l'analyse 746 on trouve un document qui n'apparaît pas dans l'autre analyse. Il concerne l'utilisation du procédé SCORIM ("Shear Controlled Orientation in Injection Moulding") pour le moulage d'un polymère à cristaux liquides avec l'objectif d'éliminer les effets néfastes des lignes internes de soudure. Or précisément les mots de "weld " (soudure) et SCORIM s'ajoutent à ceux de "injection" et "moulding" pour faire partie des mots ayant la plus forte contribution dans cette analyse.

## RESULTATS N°10

COMPARAISON ENTRE L'ANALYSE EFFECTUEE A PARTIR DES MOTS AYANT UNE FREQUENCE MINIMUM DE 1% DU NOMBRE DE DOCUMENTS LA BASE (ANALYSE 746) et ANALYSE EFFECTUEE A PARTIR DU NOMBRE D'OCCURRENCE DES MOTS (ANALYSE 673)

MOTS COMMUNS à A.746 et A.673 : *injection et moulding*

MOTS SPECIFIQUES de A.746 (absent de A.673) : *weld et SCORIM*

*Document spécifique de l'analyse 746 (absent de A.673) sur le facteur 3 négatif :*

RESUME NO 00000505

TI - ENHANCEMENT OF INTERNAL WELD LINE STRENGTH IN THERMOTROPIC LIQUID CRYSTAL POLYMER MOULDINGS

AB - Shear Controlled Orientation in Injection Moulding (Scorim) was applied to moulding of a thermotropic liquid crystal polymer with the aim of eliminating the detrimental effects of internal weld lines. The Scorim technique was shown to induce macroscopic displacement of the molten material within the mould cavity. It was found that the initial displacement of the melt, applied immediately after the cavity had been filled, was sufficient virtually to erase the weld line by re-orientation of molecules and fibre filler. The TS of the Scorim modified weld was found to be up to 16 times that of the weld in conventionally moulded bars. Application of Scorim to a moulded bar without internal weld lines resulted in a 90% improvement in strength. 30 .

Dans le cas de l'analyse 673 ce sont les mots " Ferromatik" (nom de la société ayant mis au point le procédé d'injection "Push-Pull") et "process" (procédé) qui la spécifient par rapport à l'autre. Il s'agit toujours d'injection de LCP mais avec la connotation particulière ajoutée par ces deux mots d'où la présence de ces trois documents qui ne se retrouvent pas dans l'autre analyse :

## RESULTATS N° 11

COMPARAISON ENTRE L'ANALYSE EFFECTUEE A PARTIR DES MOTS AYANT UNE FREQUENCE MINIMUM DE 1% DU NOMBRE DES DOCUMENTS DE LA BASE (ANALYSE 746 : A.746) et ANALYSE EFFECTUEE A PARTIR DU NOMBRE D'OCCURRENCE DES MOTS (ANALYSE 673 : A.673)

MOTS COMMUNS à A.746 et A.673 : *injection et moulding*

MOTS SPECIFIQUES de A.673 absents de A.746 : *Ferromatik et process*

*Documents spécifiques de l'analyse 673 (absents de A.746) sur le facteur 3 négatif :*

Document n° 891

TI - HOW TO PROCESS LCP s

AB - Improvements in processing and materials in liquid crystal polymers (LCPs) for the 1990s are discussed in detail. The challenge to make LCP's more useful as engineering materials is examined. New processing technology such as Klockner Ferromatik Desma's 'Push-Pull' LCP injection moulding process is described. The article also discusses improvements in material and mould design, new extrusion techniques for LCP sheeting/films, blending of LCP's with other resins, and the use of the Addiposit process to plate Hoechst Celanese's Vectra C-810 filled LCP. Plans for research and development of LCPs in the next few years by five suppliers are outlined and some companies who are reconsidering their commitment to LCPs are mentioned.

Document n°911 :

TI - MOULDING AND PROCESS PROPERTY RELATIONSHIPS OF A WHOLLY AROMATIC THERMOTROPIC LIQUID CRYSTAL POLYMER

AB - Three grades of Xydar copolyesters (unfilled, glass filled and mineral filled) were analysed for process/property relationships. Evaluations of the polymer physical properties as a function of mould temperature, injection speed, injection pressure and part thickness are discussed. A correlation is presented to describe the interaction of the process variables as well as a method to predict the moulded properties as a function of process parameters.

Document n°854 :

TI - TRENDS IN INJECTION MOULDING TECHNOLOGY

AB - A discussion is presented on developments in injection moulding techniques, special attention being paid to thermoplastics processing. Advances made in mould changing systems, handling of materials and mouldings as well as process and quality control are described, as are trends in liquid

crystal polymer processing, manufacture of thermoplastics structural foams and hollow bodies. Injection compression moulding is also considered. Economics of injection moulding are assessed.
--

Ces différences ne font pas apparaître une réelle perte d'information sur le plan de la thématique : les documents communs font allusion au procédé " push-pull " de Kloeckner Ferromatik par exemple et l'analyse 673 n'ajoute pas d'information sur ce plan. La différence porte plutôt sur la qualité des documents émergents, et sur des nuances d'approches : la présentation du procédé "push-pull" par le problème des lignes de soudure par exemple.

De même, au niveau de la métaclé du facteur 4 positif, les différences portent sur des nuances. Cette métaclé rassemble des articles de rhéologie : il s'agit de mesurer la pression (*pressure*), la viscosité élongationnelle (*elongational viscosity*) ou de cisaillement (*shear viscosity*) et la vitesse de cisaillement (*shear rate*) ainsi que l'effet de la pression sur la viscosité, d'un flux de matière de LCP dans un rhéomètre.

Mots ayant la plus forte contribution dans analyse 746 (classés dans l'ordre décroissant) pour le facteur 4 négatif	Mots ayant la plus forte contribution dans analyse 673 (classés dans l'ordre décroissant) pour le facteur 4 négatif
<i>Viscosity- viscosities- shear- elongational- pressure- flow- capillary- rate- rheometer.</i>	<i>Pressure- shear- flow- elongational. - viscosities- rheometer- viscosity</i>
Documents n°777- 565- 543	Documents n°565- 777- 564- 576- 794

Les documents communs aux deux analyses (numéros en gras dans le tableau) sont les documents 565 et 777.

Le document 543, qui ne se trouve que dans A.746, concerne *seulement* la mesure de la viscosité : contrairement aux documents communs aux deux analyses (n°565 et n°777), il ne contient pas le mot " pression " ; d'où le poids plus fort de ce terme dans l'analyse 746.

Dans l'analyse 673 les 3 documents (564, 576, 794) qui la spécifient par rapport à A.746 (dont ils sont absents) concernent plus particulièrement :

- la vitesse de cisaillement et la viscosité
- l'estimation de la viscosité extensionnelle et de cisaillement pour prévoir la pression dans la cavité
- la détermination de la tension interfaciale de mélanges (blends) contenant un LCP : cette étude passe en particulier par la détermination de la viscosité élongationnelle .

Malgré les différences entre les deux analyses de plus en plus notables au fur et à mesure de l'avancée dans le nombre de facteurs pris en compte, on retrouve dans l'une comme dans l'autre les grands thèmes : les applications électroniques et les circuits imprimés, les films barrière.

### III-2-2.3 Les différences avec l'analyse 296

Les thèmes sont beaucoup mieux formés que dans les autres analyses et elle est plus riche.

On constate :

- des thèmes plus développés : c'est le cas par exemple des stratifiés et des fibres de carbone
- en revanche, un thème est davantage développé dans l'analyse 673 : il s'agit des données marché.
- les mélanges ("blends") et les études thermiques ne forment pas un thème

### III-2-2.4 Conclusion

L'analyse de type A.746 basée sur les termes ayant une fréquence minimale correspondant à 1% du nombre des documents du corpus) a peu d'intérêt en elle-même, c'est-à-dire qu'elle *se prête peu à l'interprétation directe* car elle restitue de nombreux mots mais pas assez de documents ayant une qualité de représentation suffisante. Elle est en revanche très utile pour sélectionner les termes destinés à entrer dans une seconde analyse. Les mots des métaclés sont alors utilisés pour construire une nouvelle matrice mots\*documents : A.746 sert alors de base à une analyse de second niveau. Nous appellerons la première "analyse de premier niveau" et la seconde "analyse de second niveau".

Cette analyse de second niveau (de type A.296), faite à partir des mots répondant à un seuil de contribution à la valeur propre issue d'une première analyse de type A.746, est *la plus riche*. Toutefois elle fait perdre certains regroupements : celui du moulage par injection des LCP (*LCP* c'est-à-dire polymères à cristaux liquides) n'est pas caractérisé dans cette analyse même s'il ne s'agit pas nécessairement d'une perte d'information car les mêmes documents se retrouvent ailleurs mais associés différemment.

### III-2-3 ANALYSE DE DOCUMENTS REJETES

L'intérêt des méthodes d'analyse de données, notamment des diverses méthodes d'analyse factorielle, est de synthétiser de grandes quantités de données. Or toute synthèse génère une perte d'information : la mise en évidence de relations nouvelles entre des données se fait au prix de la non visibilité des informations n'entrant pas dans des relations.

Pour comprendre la nature de l'information perdue dans les analyses, qu'elle soit due à une analyse de premier niveau ou de second niveau, nous avons analysé les documents n'apparaissant pas dans les trente premiers facteurs des analyses que nous avons comparées précédemment. Il s'agit des analyses LCP (polymères à cristaux liquides).

Les résultats ne montrent pas de thèmes radicalement nouveaux mais :

- d'une part, un centrage sur le thème de la structure des matériaux et les propriétés des mélanges. Rien n'apparaît de nouveau sur les données commerciales par exemple.
- d'autre part, un approfondissement des thèmes déjà mis en évidence.

Ainsi, par exemple, l'analyse de la structure des LCP n'est plus centrée sur les techniques physiques qui permettent l'analyse mais sur les propriétés physiques (optiques, électriques, électroniques, ferroélectriques) obtenues à partir de l'organisation des chaînes de LCP<sup>49</sup>. Ces documents sont d'ailleurs associés à quelques documents sur la synthèse de nouveaux LCP. Ce thème de la synthèse, non présent dans l'analyse de premier niveau, représente toutefois très peu de documents dans toute la base (on le vérifie en interrogeant la base par équation logique).

De même le thème des mélanges constitutif d'un matériau à base de LCP est ici centré sur les composites, notamment les composites auto-renforcés ("self-reinforcing") où l'orientation macromoléculaire du LCP joue le rôle des fibres dans les composites<sup>50</sup>.

Ces résultats nous laisse penser que l'on pourrait imaginer prendre connaissance du contenu d'une base de documents par une méthode utilisant l'analyse factorielle des correspondances

<sup>49</sup> Nous citons le facteur rassemblant les documents sur ce thème dans *Résultats Détaillés, Résultats n°12 p.193-195*

<sup>50</sup> Voir *Résultats Détaillés, résultats n°13 p.196*

pour dégager tout d'abord les thèmes principaux, puis analyser les rejets et les rejets de rejets etc. L'intérêt serait ainsi de naviguer dans une base par son contenu (thèmes, sous-thèmes) plutôt qu'en tâtonnant par une sélection arbitraire de documents.

### III-2-4 LES NOTIONS DE SYNTHÈSE ET DE THEME.

Comparaison entre une synthèse issue d'un groupe d'experts et une synthèse obtenue par AFC sur les mêmes documents

Pour mieux comprendre le type de synthèse effectuée par une analyse factorielle, nous avons comparé un état de l'art rédigé par la documentaliste du Pôle Européen de Plasturgie au terme d'un travail avec le groupe d'industriels demandeurs de l'étude et celle effectuée par une analyse factorielle sur les documents sélectionnés par la documentaliste.

Le sujet était : "le recyclage du polyuréthane" pour lequel 528 documents ont été sélectionnés par la documentaliste. Ces 528 documents ont été lus par le groupe d'industriels, chacun s'étant réparti une partie des documents. Des réunions de mise en commun des résultats se sont ensuite produites et la synthèse finale a été effectuée par la documentaliste. La collecte des documents a été réalisée de façon empirique : pour chaque base de données certains documents ont été gardés d'autres rejetés selon l'intérêt qu'ils semblaient manifester *a priori*. Ils ont été sélectionnés comme dans toute étude documentaire classique, où les requêtes booléennes sur les serveurs de bases de données ont été couplés aux moyens intellectuels d'interprétation, eux-mêmes reposant sur la capacité de lecture des acteurs en présence. De plus, ces documents n'ont pas été obtenus à partir d'une seule et même requête mais tantôt avec le terme de *reclaim* tantôt avec celui de *recycling*.

Nous avons pu reprendre ces 528 documents sous le format bibliographique (format comprenant des informations sur l'identité du document et, sur le plan du contenu : le résumé et les mots-clés). Nous n'avons pas pu obtenir les textes intégraux des articles sous forme électronique. Or certains de ces articles avaient été lus par le groupe.

Une seconde difficulté s'ajoute à celle-ci : elle consiste à comparer des synthèses issues de données sélectionnées en fonction d'une connaissance préalable du sujet (les résultats de la synthèse sont, d'une certaine façon, déjà contenus dans la sélection).

Malgré ces difficultés, nous avons effectué la comparaison entre l'état de l'art et l'analyse des correspondances à partir du même corpus de documents afin de mettre en évidence les mécanismes de la synthèse dans les deux cas.

Le recyclage du polyuréthane consiste à refaire un nouveau matériau ayant de bonnes propriétés à partir de polyuréthane usagé.

Dans le cas de l'état de l'art obtenu par le travail de groupe, on obtient une typologie des recyclages associée à une définition. Ainsi définit-on les différents types de recyclage mécanique (broyage/pulvérisation, moulage par compression, compression adhésive, utilisation de déchets réduits en poudre comme charge), le recyclage chimique (pétrochimie, chimiolyse).

Dans le cas de l'analyse que nous avons effectuée, les documents sélectionnés sur les métaclés et ceux cités dans l'état de l'art étaient à peu près les mêmes mais associés différemment. Seuls les procédés de chimiolyse se sont distingués en tant que thème de la même manière dans l'état de l'art et dans nos analyses. L'état de l'art ne s'est pas caractérisé pas par le fait qu'il contenait des informations nouvelles et singulières que l'analyse statistique aurait occultées mais bien par le type de synthèse qu'il effectuait.

Dans l'analyse statistique ce ne sont pas les types de recyclage qui sont mis en évidence mais des regroupements de documents autour, par exemple, du rôle du phosphate, participant tantôt d'un liant, d'un plastifiant ou d'un agent supprimant les fumées (les fumées issues de la combustion). Ce sont les conditions expérimentales du recyclage qui créent aussi des regroupements sur un même facteur : ainsi les termes de degrés, bar, minutes sont associés avec une forte contribution à la valeur propre et mettent en évidence un ensemble de brevets.

On ne peut considérer qu'une synthèse (état de l'art ou étude statistique) est plus ou moins générale qu'une autre, qu'une analyse statistique ne mettrait en évidence que le général alors qu'un état de l'art serait plus sensible à la singularité. La réalité est plus complexe.

S'il est vrai qu'une analyse statistique ne peut mettre en évidence que des relations et donc occulte nécessairement le singulier, elle ne se cantonne pas néanmoins dans la généralité. Un facteur correspond à un groupe d'éléments qui se discriminent par rapport à l'ensemble. Le point commun, la tendance s'exprime au centre de gravité. Ce groupe d'éléments assemblés sur une partie du facteur pourrait être considéré comme une tendance locale.

Une étude statistique ne permet pas de prendre connaissance d'un sujet comme le fait un état de l'art effectué par un documentaliste ou un groupe de travail. Son intérêt est seulement de faire apparaître des informations issues d'associations non décelables à la lecture, c'est-à-dire essentiellement selon une approche séquentielle des contenus des textes.

### III-2-5 CIBLAGE D'UNE EVOLUTION CHRONOLOGIQUE

Nous avons voulu tester l'intérêt d'une analyse croisant les mots des documents et les années afin de voir dans quelle mesure ce type d'analyse pouvait nous mener à repérer une chronologie des thèmes.

Ce test a été fait à partir de l'analyse *CAPTEURS* (n°2.2; cf p.210). Il a été conçu selon la stratégie suivante :

Une première analyse a été réalisée à partir du tableau mots\*années en regroupant les années 1980 à 1982 à cause du faible nombre de mots appartenant à chacune d'entre elle (analyse n°1). Au terme de cette première analyse, est apparu un regroupement entre les années : c'est celui que l'on observe dans le graphique de résultats n°14 (p.159). Les périodes 1980-84, 1985-88, 1989-94, 1995-97 sont des résultats de l'analyse : les points "années"les représentant se sont associés selon ce découpage.

Une deuxième analyse (analyse n°2) a été alors effectuée à partir du tableau mots\*périodes (celles obtenues par la précédente analyse).

Nous avons ensuite procéder à une classification des mots d'après les coordonnées de la précédente AFC (analyse n°2) et effectuer une troisième AFC (analyse n°3) croisant les classes de mots (de c1 à c4 sur le graphique p.159) et les périodes. C'est cette troisième analyse (analyse n°1) qui est représentée dans le graphique ci-après [Résultats n°14 (p.159)]. Elle est représentée avec les mots placés en éléments supplémentaires (ils ont été ajoutés après le calcul des facteurs).

Ces résultats illustrent surtout la politique de sélection de la base de données RAPRA, dont nous avons utilisé les données dans cette expérience c'est-à-dire la manière dont elle a sélectionné l'information sur les capteurs.. La période 1995-97 se caractérise par les mots

illustrant le thème "biocapteurs" : cela correspond aux résultats des analyses factorielles issues de la mise à jour du sujet (analyses 2) qui se spécifiaient par rapport à l'ancienne (analyse 3) en ce qu'elles introduisaient ce sujet qui apparaissait dès le premier facteur.

La période 1989-1994 illustre exactement les thèmes centraux pour la plasturgie alors que les documents de la période 1995-1997 sont en dehors du domaine puisqu'il touche aux polymères utilisés comme capteurs. La base des documents rassemblés sur ce sujet suggère donc un manque d'avancée significative sur le thème de la relation entre capteurs et contrôle de procédés qui est au cœur du sujet.

L'intérêt des résultats de cette démarche nous semble résider autant dans la forme du graphique obtenu que dans l'interprétation. En effet, pourrait-on caractériser la forme que peut prendre la discrimination des périodes en fonction du vocabulaire, selon les bases et les sujets traités ? Imaginons que la forme du graphique soit une "boule", pourrait-on en déduire que le sujet représenté par la base de documents est un domaine qui stagne ? Nous n'avons pu dans le cadre de ce travail tester ce genre d'hypothèse qui aurait permis de typer les formes de graphiques obtenues et leur signification.

### RESULTATS N°14

#### EVOLUTION CHRONOLOGIQUE DU VOCABULAIRE DE LA BASE RAPRA SUR LE SUJET *CAPTEURS*

Représentation sans les points double.

**C1, c2, c3, c4** sont des classes de mots : leur contenu est détaillé dans *Résultats Détaillés,*

*résultats n°14 suite p.197-198.*

AXE HORIZONTAL(f\_\_1)--AXE VERTICAL(f\_\_2)

NOMBRE DE POINTS : 110

==ECHELLE : 4 CARACTERE(S) = 0.13 1 LIGNE = 0.05

DEPENDANT			1
CONTAINERS			2
SITU			3
AUTOMATED HDPE ELECTROMAGNETIC			4
FLOWING DYNISCO			5
CORRELATED AUTOCLAVE RTM SPC			6
RHEOLOGICAL LASER			7
GRAPHITE ENGINEERING CURE REACTION VISCOSITY			8
COMPOSITES REVIEWED			9
FLUORESCENCE SHOWN TG			10
COMPOSITE REPORTED <a href="#">C2 C3 1989-94</a> ASSURANCE			11
EMPHASIS			12
MATRIX THIN			13
INC CHEMICAL METAL			14
POLYPYRROLE SO <a href="#">C1 1995-97</a>			15
INDICATOR POTENTIAL STRAIN			16
ENZYME ELECTRODE LIQUID			17
FIXED SILICONE HEATER <a href="#">1985-1988</a>			18
GLUCOSE RAM			19
SUBSTRATE			20
FOAM			21
ORIENTATION VISION			22
CONSISTING PRODUCTIVITY			23
ALLOW			24
CHEMICALS INCORPORATED BUILT <a href="#">C4 1980-84</a>			25
BETA CONSISTS			26
PRE INFRA			27
PATTERN			28
ANGLE SCANNER			29
PROTECTION			30
			31
			32
			33
			34
			35
			36
			37
			38
			39
			40
			41
			42
			43
			44
			45
			46
			47
			48

### **III-3 DISCUSSION DES RESULTATS**

#### III-3-1 La désambiguïisation sans dictionnaire

Nous avons constaté que l'analyse factorielle des correspondance permettait de désambiguïiser des thèmes sans recours à des techniques linguistiques reposant sur l'usage de dictionnaire.

Ainsi avons-nous pu constater que, pour le sujet "Capteurs et transformation des polymères", les polymères utilisés comme capteurs et les capteurs utilisés dans la transformation des polymères, représentant une nuance impossible à désambiguïiser au niveau de l'interrogation de la base de données, se discriminaient en revanche parfaitement bien avec l'analyse factorielle. Cette différence est vérifiable dans l'ensemble des analyses *capteurs* effectuées. La désambiguïisation a donc été stable. Nous reproduisons l'ensemble des résultats dans : *Résultats Détaillés Résultats n° 15*, p. 199-207.

#### III-3-2 La lemmatisation rendue caduque

Le chapitre III-2-1 a montré que l'absence de lemmatisation ne perturbait pas la cohérence générale des résultats. On peut considérer qu'il s'agit là d'un avantage important quand on sait qu'elle n'est pas seulement lourde à gérer (elle suppose la constitution d'un lexique et d'une grammaire), mais qu'elle est également réductrice de sens y compris dans le cas des textes scientifiques que l'on ramène trop souvent à un langage pauvre et sans nuances.

Pour illustrer ce point, nous proposons d'examiner le cas de l'association de la forme plurielle et singulière d'une forme alors qu'elles sont traitées comme deux formes distinctes dans le tableau de contingence soumis à l'AFC.

**Dans la majeure partie des cas la forme singulière et la forme plurielle d'un terme sont associées sur le même facteur.** Nous le vérifions dans l'ensemble des analyses. Toutefois il peut arriver que seule l'une des formes (en général il s'agit du singulier) apparaisse sur un facteur. Ce cas de figure n'est pas sans signification.

Ainsi, par exemple, sur les 9 occurrences du terme *membrane* dans les facteurs d'une analyse portant sur les capteurs, on trouvera 8 fois l'association de *membrane* et *membranes* et une seule fois le terme de *membrane* seul.

Ce dernier cas de figure se produit sur le facteur 6 de l'analyse CAPTEURS (analyse 2-2) que nous reproduisons ci-après.

## RESULTATS N°16

**PRESENCE DU MOT MEMBRANE AU SINGULIER SEULEMENT, DANS L'ANALYSE FACTORIELLE CAPTEURS 2-2**

POUR LE FACTEUR 6, LISTE DES MOTS A COORDONNEE NEGATIVE, IL Y A 28 MOTS EXPLICATIFS :

SEUILS POUR LA CONTRIBUTION 2\*MOY = 5.29 3\*MOY = 7.94

LIGNE	FACTEUR	QUALITE	CONTRI		
23	-0.50	46.22	12.02	163	INJECTION
23	-0.42	19.04	5.75	212	MOULD
23	-0.51	63.24	15.87	213	MOULDING
23	-0.59	32.29	9.63	284	RESIN
24	-0.77	16.24	6.85	79	DIELECTRIC
24	-0.80	28.43	6.59	267	PRODUCTION
24	-0.66	19.17	6.60	295	RUBBER
25	-0.88	50.12	19.43	67	CURE
25	-0.88	19.07	5.67	195	MEMBRANE
26	-1.29	43.85	11.63	255	POLYMERISATION
27	-1.63	18.05	6.04	21	BIOSENSORS
27	-1.40	39.80	10.01	94	ELECTRODE
27	-1.40	38.26	13.35	120	FLUORESCENCE
28	-1.87	12.72	5.47	103	ETHYLENE
28	-1.83	43.85	10.48	155	IMMOBILISATION
28	-1.90	6.99	5.63	200	MICHELIN
29	-2.14	58.44	18.85	100	ENZYME
29	-2.08	22.89	10.49	246	PLATEN
31	-2.49	61.32	15.06	9	AMPEROMETRIC
31	-2.46	86.19	37.19	131	GLUCOSE
31	-2.54	52.93	10.11	136	GROWTH
31	-2.64	71.18	13.90	199	METHYL
31	-2.51	46.46	11.64	363	UV
32	-2.92	80.90	18.21	198	METHACRYLATE
33	-3.12	46.47	15.27	61	COPOLYMERISATION
33	-2.99	98.29	26.76	228	OXIDASE
35	-3.64	76.69	26.44	330	STYRENE
50	-7.54	108.14	88.96	177	LATEX

QUALITE = 67 ( 7 ) documents

Liste des titres des 7 documents associés:

**Résumé Document 00001197 ==> Ligne 29 Fac = -2.12 Qualité = 77.81 Contribution = 22.42**  
**@TI ELECTROCHEMICAL FABRICATION OF AMPEROMETRIC GLUCOSE ENZYME ELECTRODE BY IMMOBILISING GLUCOSE OXIDASE IN ELECTROPOLYMERISED POLY(3,3'-DIAMINOBENZIDINE)(I) FILM ON PALLADINISED GLASSY CARBON ELECTRODE**  
**@SO Analyst; 121, No.7, July 1996, p.971-5**

**Résumé Document 00001305 ==> Ligne 29 Fac = -1.96 Qualité = 71.67 Contribution = 8.79**  
**@TI STRUCTURE AND PROPERTIES OF THE COMPOSITE MEMBRANE OF REGENERATED SILK FIBROIN AND PVAL AND ITS APPLICATION TO AMPEROMETRIC TETRATHIAFULVALENE-MEDIATING GLUCOSE SENSOR**  
**@SO J.Macromol.Sci.A; A33, No.2, 1996, p.209-19**

**Résumé Document 00001022 ==> Ligne 30 Fac = -2.43 Qualité = 70.91 Contribution = 18.43**  
**@TI AMPEROMETRIC BIOSENSORS FOR DETECTION OF CHOLINE AND LACTATE BY CHOLINE OXIDASE AND LACTATE OXIDASE IMMOBILISATION IN MICROPOROUS MEMBRANES**  
**@SO Espoo, 1996, p.39. 30cms. 10/2/97**

**Résumé Document 00001196 ==> Ligne 30 Fac = -2.37 Qualité = 78.55 Contribution = 17.53**  
**@TI INTERFERING FACTOR IN THE GLUCOSE OXIDASE SENSING SYSTEM WITH POLYPYRROLE/GLUCOSE OXIDASE MEMBRANE**  
**@SO Synthetic Metals; 71, Nos.1-3, 1st April 1995, p.2245-6**

## RESULTATS N°16 (fin)

PRESENCE DU MOT MEMBRANE AU SINGULIER SEULEMENT, DANS L'ANALYSE FACTORIELLE  
CAPTEURS 2-2

**Résumé** Document 00001209 ==> Ligne 32 Fac = -2.73 Qualité = 97.92 Contribution = 10.59  
@TI ELECTROCHEMICAL SYNTHESIS AND SENSOR APPLICATION OF  
POLYDIAMINOBENZENE  
@SO Synthetic Metals; 79, No.1, 30th April 1996, p.57-61

**Résumé** Document 00001444 ==> Ligne 34 Fac = -3.27 Qualité = 94.31 Contribution = 47.33  
@TI UV AND FLUORESCENCE CHARACTERISATION OF STYRENE AND METHYL  
METHACRYLATE POLYMERISATION  
@SO J.Appl.Polym.Sci.; 57, No.3, 18th July 1995, p.363-70

Ces sept documents, dont seuls les mots du résumé ont été traités, peuvent contenir le terme de *membranes* même si sur le facteur c'est *membrane* qui est sélectionné. Cette sélection n'est pas anodine.

En effet, ici la présence du singulier signifie que le terme a un poids sémantique particulier en tant que singulier. Les 7 documents de ce facteur ont pour point commun de traiter de l'utilisation de polymères conducteurs comme capteurs de certaines réactions électrochimiques comme l'oxydation du glucose. Le but est de mesurer la concentration d'un produit biologique, typiquement le glucose. L'utilisation d'une membrane entre deux électrodes est un procédé pour provoquer un mouvement d'électrons dû à la réaction d'oxydoréduction. Ce mouvement permet la mesure de la concentration du produit. La membrane joue donc un rôle en elle-même et sa nature est toujours spécifique : elle est donc le plus souvent désignée au singulier. Plus précisément, le singulier a deux fonctions : d'une part, il signale qu'il est fait référence à une membrane bien particulière ("*a PVDF affinity membrane (Immobilon), an amphoteric nylon 66 membrane (Biodyne)etc.*") ; d'autre part, que la membrane est un sujet central repris plusieurs fois, d'où la présence de l'article défini : "*the track-etch membrane, the polypyrrole/glucose oxidase membrane, the membrane*".

Nous citons les fiches bibliographiques des documents de ce facteur ayant le terme de *membrane*.

## RESULTATS N°17

Sens de la métaclé avec seulement *MEMBRANE* au singulier  
FICHES BIBLIOGRAPHIQUES DES DOCUMENTS DU FACTEUR AYANT LE TERME DE *MEMBRANE*  
SEUL

Résumé N0 00001022

@**TITRE** : AMPEROMETRIC BIOSENSORS FOR DETECTION OF CHOLINE AND LACTATE BY CHOLINE OXIDASE AND LACTATE OXIDASE IMMOBILISATION IN MICROPOROUS MEMBRANES

@**ABSTRACT** : This report deals with the construction of amperometric biosensors for the detection of choline and lactate, using choline oxidase and lactate oxidase as the selective elements. The enzymes were immobilised in a variety of microporous membrane supports using different immobilisation protocols. Track-etch polyester or polycarbonate membranes (Nuclepore and Cyclopore) were used in the comparison, as were a PVDF affinity membrane (Immobilon) and an amphoteric nylon 66 membrane (Biodyne). The electrochemical sensor response of the membranes was compared with a spectrophotometric determination of enzyme activity for the various immobilisation methods used. The track-etch membrane appeared to be the best material for immobilisation, giving a fast response and good reproducibility.

@**DESCRIPTEURS** : AFFINITY ; AMPEROMETRIC ; BIOSENSOR ; CARBONATE POLYMER ; COMPANIES ; COMPANY ; DATA ; ELECTROCHEMICAL ; ENZYME IMMOBILISATION ; ENZYME IMMOBILIZATION ; GRAPH ; LACTATE OXIDASE ; MEMBRANE ; MICROPOROUS ; NYLON-6,6 ; PLASTIC ; POLYCARBONATE ; POLYESTER FILM ; POLYVINYLIDENE FLUORIDE ; POROUS ; PVDF ; SPECTROPHOTOMETRY ; TABLES ; THEORY ; THERMOPLASTIC ; TRACK MEMBRANE ; VINYLIDENE FLUORIDE POLYMER

RESUME N0 00001196

@**TITRE** : INTERFERING FACTOR IN THE GLUCOSE OXIDASE SENSING SYSTEM WITH POLYPYRROLE/GLUCOSE OXIDASE MEMBRANE

@**ABSTRACT** : A glucose sensor was prepared by immobilising glucose oxidase in a polypyrrole membrane during the electrochemical polymerisation of pyrrole. Glucose detection was carried out amperometrically by applying a constant potential to the polypyrrole/glucose oxidase membrane. The electrolysis of hydrogen peroxide stemming from the enzyme reaction was found to influence the electrical conductivity in the membrane.

@**DESCRIPTEURS** : AMPEROMETRY ; COMPANY ; CYCLIC VOLTAMMETRY ; DATA ; ELECTRICAL ; CONDUCTIVITY ; ELECTRICAL PROPERTIES ; ELECTROCHEMICAL POLYMERISATION ; ELECTROINITIATED POLYMERISATION ; ELECTROLYSIS ; ENZYME IMMOBILISATION ; GRAPH ; MEASUREMENT ; MEMBRANE ; PLASTIC ; POLYPYRROLE ; REACTION ; SENSOR ; TECHNICAL ; THERMOPLASTIC

## RESULTATS N° 18

### ASSOCIATION DU SINGULIER ET DU PLURIEL DE LA FORME *MEMBRANE*

Le facteur 1 de l'analyse *CAPTEURS*

IL Y A 35 MOTS EXPLICATIFS :

POUR LE FACTEUR 1 LISTE DES MOTS A COORDONNEE POSITIVE

SEUILS POUR LA CONTRIBUTION 2\*MOY = 5.29 3\*MOY = 7.94

LIGNE	FACTEUR	QUALITE	CONTRI		
1	3.69	149.69	28.27	228	OXIDASE
3	3.45	116.96	19.93	9	AMPEROMETRIC
6	3.20	145.29	43.49	131	GLUCOSE
8	2.94	58.94	13.68	21	BIOSENSORS
8	2.99	114.13	25.54	100	ENZYME
8	2.93	112.18	18.60	155	IMMOBILISATION
9	2.79	158.06	27.58	94	ELECTRODE
10	2.70	38.10	8.62	258	POLYTHIOPHENE
11	2.61	58.97	12.75	282	REDOX
12	2.48	150.43	31.04	195	MEMBRANE
12	2.48	120.58	24.34	196	MEMBRANES
12	2.53	26.65	7.57	259	POROUS
13	2.42	21.76	9.24	172	IONS
13	2.42	93.56	26.66	257	POLYPYRROLE
14	2.28	108.57	16.48	95	ELECTRODES
15	2.18	45.10	7.04	198	METHACRYLATE
16	2.07	82.72	19.03	171	ION
16	2.08	54.38	13.26	236	PH
17	1.96	38.29	8.34	1	ACID
17	2.01	53.52	18.02	229	OXYGEN
17	2.03	28.00	7.73	241	PLASMA
17	1.96	41.82	10.97	250	POLY
17	2.01	91.12	16.78	304	SELECTIVE
18	1.91	35.37	6.83	129	GEL
18	1.90	28.58	6.78	234	PERMEABILITY
20	1.64	58.03	13.32	49	CONDUCTING
21	1.52	61.24	9.58	226	ORGANIC
21	1.53	41.27	8.29	366	VAPOUR
23	1.33	82.87	16.44	116	FILMS
24	1.26	18.13	5.94	150	HUMIDITY
24	1.23	78.78	17.17	256	POLYMERS
25	1.08	25.72	5.69	2	ACOUSTIC
25	1.10	58.43	9.63	34	CHEMICAL
25	1.14	34.43	6.33	255	POLYMERISATION
27	0.92	128.10	23.27	254	POLYMER

Qualité = 67 ( 27 ) documents

En revanche, dans le facteur 1 par exemple de cette même analyse (l'analyse *CAPTEURS*), le singulier et le pluriel sont associés. Ici les documents<sup>51</sup> (facteur 1) sont inclus dans un ensemble plus large touchant aux polymères conducteurs utilisés comme biocapteurs. On trouve associées la forme singulière et la forme plurielle du terme membrane car le terme n'a pas de poids spécifique en tant que singulier. Les textes évoquant le rôle de la membrane sur ce facteur ne sont pas nécessairement centrés sur la construction de biocapteurs mais concernent plutôt, comme les textes suivants, *les matériaux pouvant entrer dans la conception de membranes* dont l'application est l'utilisation comme biocapteur :

@TI FILLED FLUROSILICONE AS MATRIX MATERIAL FOR ION-SELECTIVE MEMBRANES

@AB Thin film and containment sensors with fluorosilicone as the matrix material for **the ion-selective membrane** were studied. Potassium-, sodium- and calcium-selective sensors, with response characteristics similar to those of a conventional ion-selective electrode using PVC-matrix **membranes**, were obtained. The main features of **the new membrane** were good adhesion and improved mechanical strength compared with PVC **matrix membranes**. Furthermore, it had an optimal conductivity and a higher relative permittivity than 'conventional' silicones. 7 .

@TI POLYMER MEMBRANES IN CLINICAL SENSOR APPLICATIONS. II. DESIGN AND FABRICATION OF PERMSELECTIVE HYDROGELS FOR ELECTROCHEMICAL DEVICES

@AB The design of hydrogel membranes for ion selective sensor applications was investigated by ion permeability and coated wire ion selective electrode studies. The potential use of photolithographic techniques in thin film processing of these materials was also studied as a basis **for multiple membrane application** at the wafer stage of ion selective field effect transistor fabrication. Materials examined included plasticised PVC, polyhydroxyethyl methacrylate and hydroxyethyl methacrylate-methyl methacrylate copolymer. 52

@TI PHOTOCURABLE POLYMER MATRICES FOR POTASSIUM-SENSITIVE ION-SELECTIVE ELECTRODE MEMBRANES

@AB Photocurable oligomers based on urethane and Bisphenol A (epoxy resin) diacrylates were studied as an alternative polymer matrix for **ion-selective membranes** of ion sensors. The compatibility of the polymers with various plasticisers commonly used in **PVC-based ion-selective membrane** formulations was tested, and it was shown that **the urethane-based membranes** may hold more than 50% of a plasticiser, while epoxy-based compositions are not fully compatible with them. Potassium-sensitive electrodes with an inner reference solution and **a membrane** formed by the photo-cured urethane containing valinomycin as the ionophore were investigated. The influence of the amount and the type of photoinitiator, crosslinking agent, and plasticiser on ion response of the ion-selective electrode was studied. The developed polymer matrix was used to make an ISFET-based potassium sensor. 36 .

Ces textes contiennent presque autant d'emplois du terme *membrane* au pluriel qu'au singulier. En effet le thème du choix des matériaux pour la conception de membranes amène à évoquer celles-ci de façon génériques sous la forme "les membranes". Il n'y a pas ici de description d'une expérience pour l'utilisation de biocapteurs où la membrane joue un rôle actif spécifique.

<sup>51</sup> Les 27 documents du facteur 1 mentionnés p.164.

### **III-4 DIFFUSION DES RESULTATS. Quel système documentaire pour l'intelligence économique ?**

La synthèse des documents obtenue par AFC a pour résultat :

- 1- de mettre à jour une typologie thématique de la base de documents
- 2- de sélectionner 15 à 20% des documents pour illustrer cette typologie thématique sur les trente premiers facteurs.

Notre approche du système documentaire destiné à l'intelligence économique consiste à mettre en œuvre de façon complémentaire cette synthèse obtenue par AFC et un moteur de recherche booléen classique. La première a pour objectif de donner un guide de lecture à l'utilisateur quand le second a pour intérêt de restituer l'ensemble exhaustif des documents correspondant à un thème mis en évidence par l'AFC.

L'ensemble est accessible sur un serveur HTTP (Hypertext Transfer Protocol).

Figure 11 nous représentons les 4 modes de recherche permis par le système :

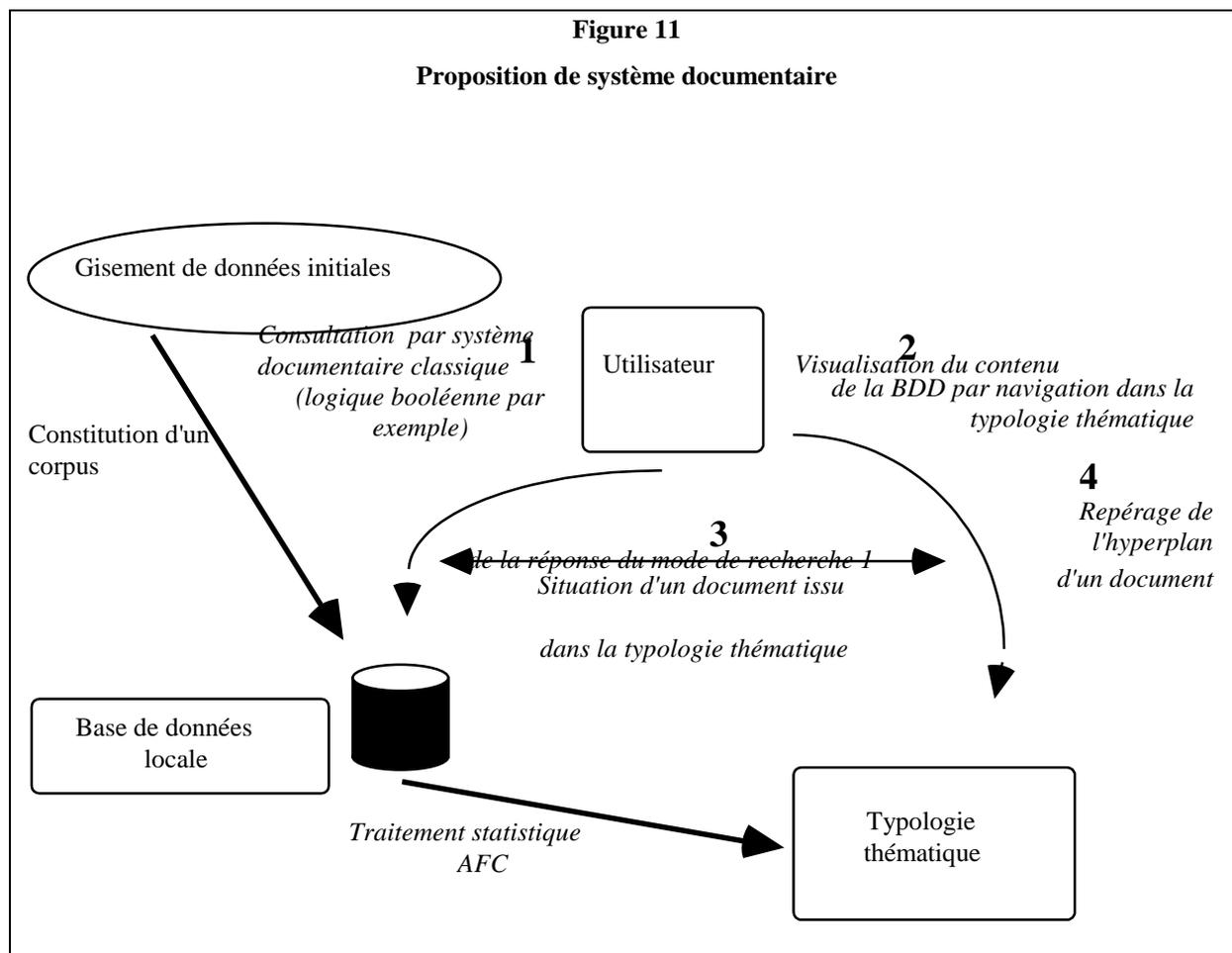
- 1- le corpus de travail contenant tous les documents concernant le sujet d'étude que nous avons appelé "base de données locale" est consultable par un moteur de recherche classique (équations logiques). Toutes les bases concernant tous les sujets traités par l'utilisateur sont capitalisées et interrogeables de cette manière. C'est le mode de consultation n°1.
- 2- les résultats des analyses factorielles sont visualisables sur le même serveur. On peut entreprendre la connaissance du sujet d'étude en navigant à l'intérieur de la typologie thématique représentée par les différents métaclés : c'est le mode de consultation n°2. Sur chaque métaclé les associations de mots et documents permettent d'obtenir cette connaissance. Le simple accès à des documents (mode n°1) permet de chercher un document ou mot particulier; le mode n°2 nous donne un aperçu de la façon dont le sujet est traité et nous invite à une première lecture guidée des documents.

- 3- Grâce au mode n°3, il est possible de passer des résultats d'une équation logique donnés par le moteur de recherche (mode 1) à leur situation dans la typologie thématique. Ces résultats sont des documents dont on peut obtenir la position sur l'un des facteurs de l'analyse. Ainsi le document apparaît "en contexte" sur une métaclé.
- 4- Le mode 4 permet de naviguer à l'intérieur de l'hyperplan d'un document <sup>52</sup>, c'est-à-dire de l'ensemble des métaclés sur lesquelles il est éventuellement présent. Si le document est multithématique, il pourra être interprété selon les différents axes d'interprétations suggérés par les métaclés.

Les calculs (les différentes analyses factorielles) ne font pas sur le serveur. Les résultats y sont ensuite transférés pour être consultés. Seules les interrogations par équations logiques se font directement en ligne. Sur le serveur sont regroupés les fichiers HTML comprenant les bases, les index, les résultats des diverses analyses soient, pour nos études 2.3 Go (Gigaoctets), 12827 fiches et 63 répertoires. Une base de 0.5 Mo a généré 5 Mo d'index afin de permettre une interrogation à deux niveaux (au niveau du paragraphe et au niveau du document).

---

<sup>52</sup> Voir p.140 sur cette notion d'hyperplan. Un hyperplan est l'ensemble des dimensions d'un document dans une analyse factorielle des correspondances mots\*documents. Ces dimensions sont les facteurs de l'analyses sur lesquels un document peut se trouver s'il est "multi-thématique" en particulier.



### III-5 CONCLUSION

Ces expérimentations ont permis d'envisager ce que peut être un travail de traitement systématique du contenu des documents susceptible de répondre aux finalités de l'intelligence économique. Elles ont été conçues et réalisées afin de savoir dans quelle mesure un tel travail pouvait être mené "en routine" sur des corpus de dimension "réelle" (correspondant à des sujets courants), appliqués à domaine industriel pluridisciplinaire et moins spécialisé que peuvent l'être des domaines scientifiques de la recherche fondamentale. Or les expériences en traitement du contenu des documents se font le plus souvent sur des micro-domaines scientifiques, sans qu'il soit possible d'évaluer une généralisation de la méthode.

De plus, nous avons constaté que relativement peu de travaux sont menés sur l'utilisation de l'analyse factorielle des correspondances en documentation (sur le texte des documents). Très utilisée en scientométrie, elle l'est également sur les textes dans le cadre essentiellement des

questionnaires d'enquêtes (réponses aux questions ouvertes), parfois des textes littéraires. Or les documents ne sont réductibles ni aux premiers ni aux seconds : la richesse du vocabulaire employé est intermédiaire entre celle des discours transcrits (fortes répétitions) et des textes littéraires (grande diversité). Les traitements intéressants pour l'interprétation ne sont donc pas tout à fait les mêmes. C'est pourquoi il nous semblait utile de contribuer à une réflexion sur l'utilisation de l'analyse factorielle en documentation.

La démarche que nous avons adoptée préconise un certain type de travail documentaire où l'expert-demandeur d'études est actif dans l'interprétation. Or tout l'effort des logiciels documentaires est d'économiser l'effort de l'utilisateur. Aboutissant à une sélection de documents classés par ordre de pertinence, l'utilisateur n'a qu'à contempler ce que le système a calculé pour lui.

Mais la sollicitation de l'utilisateur a aussi des limites : il s'attend à un certain nombre de résultats qui bloquent tout esprit de découverte. Il a tendance à juger les résultats en se rendant prisonnier d'une alternative : trivialité/bruit. Or la pertinence n'est pas pour nous un indice (une donnée quantitative) mais une organisation des connaissances contenues dans les documents.

## Conclusion générale

Confrontés à la question de savoir quels outils devaient être mis en place pour développer un programme d'intelligence économique au service d'une filière professionnelle (en l'occurrence : la plasturgie) nous nous sommes aperçus que la notion même d'intelligence économique n'avait pas été élucidée. Au-delà d'une définition générale qui pourrait être "système de renseignement sur l'environnement concurrentiel d'un acteur économique en vue de la défense de ses intérêts" personne ne semblait s'être intéressé à la question de savoir pourquoi cette notion est aujourd'hui subitement d'actualité depuis le début des années 1990 en France. L'intelligence économique n'est pas un concept nouveau, comme nous l'avons montré au début de notre étude : une théorie du rôle des systèmes de renseignement dans la gestion était à l'ordre du jour dès la fin des années 1960 aux Etats-Unis en même temps que se développaient d'extraordinaires investissements dans les technologies de l'information. L'actualité de l'intelligence économique est liée à l'idée que ces systèmes de renseignement doivent jouer un rôle offensif dans ce qu'il est convenu d'appeler "la guerre économique". Le renseignement dont nous parlons ici n'est pas le service questions-réponses que nous avons coutume d'appeler "renseignement" en français. Il s'agit d'une gestion militaire de l'information et de la communication : les oppositions entre espionnage et intelligence économique le rappellent sans cesse. Le concurrent remplace l'ennemi politique. Or la notion d'intelligence économique précède celle de guerre économique. Nous avons donc dû comprendre comment cette notion avait évolué et comment elle avait été importée en France.

L'intérêt porté aux systèmes documentaires dans cette perspective découle à la fois du changement de statut du document, dans l'entreprise notamment, et des possibilités offertes par le traitement massif des données. Le document n'est plus seulement une "référence" conçue pour être retrouvée : il représente davantage qu'une archive ; il est aussi une parcelle de connaissance que le traitement et la gestion des documents doit capitaliser et mettre en relation. La gestion des connaissances se substitue à la recherche d'information. Elle est, à ce titre, conçue comme une ressource stratégique de l'entreprise au même titre que l'exploitation d'autres gisements de données.

Les méthodes de traitement des documents relèvent soit de l'informatique documentaire quand il s'agit de retrouver les documents, soit de la bibliométrie quand il s'agit d'en extraire des informations de synthèse. Toutefois la bibliométrie s'est traditionnellement cantonnée à

l'exploitation des informations secondaires des documents (les informations relatives à son identité). Il convient donc de définir un autre champ, que nous avons appelé "analyse de contenu", qui partage les méthodes et les objectifs de la bibliométrie (méthodes quantitatives pour mettre en évidence des informations de synthèse), mais dont l'objet est le texte des documents. Dans ce domaine, qui relève de l'analyse de données textuelles, les travaux ont essentiellement porté sur le traitement des enquêtes (réponses aux questions ouvertes), des entretiens ou discours transcrits et, moins souvent, des textes littéraires. Quant à ceux qui traitent le texte dans le cadre de la documentation, ils ne prennent en compte, en général, que de faibles échantillons de mots ou de textes et surtout, procèdent rarement à une évaluation des méthodes utilisées. Nous avons donc souhaité, dans le cadre de notre travail, tester une méthode sur des corpus de taille relativement importante et des échantillons de mots sélectionnés de manière objective. A partir de ces données, nous avons voulu, non pas nous contenter de montrer que l'application d'une méthode (en l'occurrence l'Analyse Factorielle des Correspondances) permettait d'obtenir des résultats cohérents, mais nous avons voulu comprendre à quelles conditions elle le permettait et pour quelle qualité de résultats. Or ce travail ne nous semblait pas avoir été effectué, l'Analyse Factorielle des Correspondances étant très souvent utilisée à titre illustratif à travers l'interprétation des premiers facteurs, et non comme méthode documentaire méritant, dans ce cadre, d'être évaluée. Nos expérimentations ont poursuivi cet objectif. Il était, en effet, essentiel pour nous, de pouvoir évaluer la qualité de l'information obtenue afin de savoir dans quelle mesure elle pouvait être utilisée comme une aide à la décision dans le cadre de l'intelligence économique.

Ces expérimentations ont montré comment il était possible de traiter de grandes matrices creuses de données linguistiques par analyse factorielle des correspondances. L'intérêt de cette démarche réside, d'une part, dans l'indépendance vis à vis de ressources lexicales et grammaticales préconstituées (absence de procédures lourdes de lemmatisation, de désambiguïsation, d'analyse syntaxique pour la délimitation d'unités linguistiques complexes), indépendance qui s'accompagne d'une simplicité pratique. L'intérêt réside, d'autre part, dans la richesse des aspects linguistiques pris en compte, aspects que l'on a souvent tendance à négliger quand il s'agit de textes scientifiques.

Nous avons testé la cohérence, la stabilité et la perte d'information issues du traitement par AFC de telles matrices de données brutes. Ces tests nous ont permis de mettre en évidence des stratégies intéressantes : la sélection des mots en fonction d'un critère de contribution à la

valeur propre pour effectuer une seconde AFC à partir de laquelle se fera l'interprétation, les analyses successives de documents rejetés par une analyse précédente, une certaine stratégie pour repérer une évolution chronologique du contenu des documents.

Nous avons proposé une conception de l'interprétation basée sur la notion de métaclé, comprise comme association de mots et de documents sur le pôle positif ou négatif de chaque facteur. Chaque pôle est ainsi envisagé pour lui même comme manifestation d'un thème, interprétable à partir d'un ensemble de mots utilisés comme guide de lecture d'un ensemble de documents.

A ce mode d'interprétation nous avons associé un système documentaire permettant à un utilisateur-expert de naviguer au sein de "son sujet", comprenant à la fois une base de documents interrogeables par équations logiques et l'analyse de cette base par métaclés. A cette analyse peuvent s'ajouter des cartes factorielles bi-dimensionnelles destinées, soit à élucider des thématiques, soit à repérer une évolution chronologique. Ce type de recherche documentaire est destiné à l'expertise en général, c'est-à-dire à des spécialistes recherchant des informations sur leur domaine d'expertise ; il n'est pas conçu pour un public hétérogène recherchant de l'information à partir d'un corpus hétérogène. L'intelligence économique suppose la mise en place de tels systèmes documentaires spécialisés. Nous avons donc proposé une démarche où l'utilisateur est actif et travaille à l'interprétation des résultats au lieu de prendre connaissance passivement d'une synthèse élaborée par un tiers. Le système a seulement pour objectif de le guider dans sa lecture personnelle de la base de documents concernant son sujet.

L'enjeu du traitement de l'information documentaire pour le milieu économique est aujourd'hui très important.

L'aspect lié à la surveillance de la concurrence, à la version légale de l'espionnage économique, comme le rappellent souvent les américains, n'est qu'un aspect de l'intelligence économique. L'information dans l'économie n'est pas seulement une arme pour se défendre ou attaquer, c'est aussi une valeur économique. L'engouement actuel pour l'intelligence économique pourrait apparaître de ce point de vue comme un mode de participation au marché de l'information.

Le poids relatif de la partie immatérielle dans la valeur ajoutée ne cesse d'augmenter car, de plus en plus, le contenu informationnel dans un produit dépasse, en valeur, son contenu en énergie, en matière première et en heures de travail manufacturier.

Nous reprenons la définition que Yolin [1998] donne des "coûts informationnels" (coûts relatifs au contenu informationnel du produit):

*"Nous entendons par coûts informationnels (les):*

- *Etudes préliminaires de marché : de quoi le client a besoin*
- *Coûts de conception: R&D, mise au point, protection juridique...*
- *Coûts de fabrication de la partie immatérielle: formation des opérateurs, élaboration des modes d'emploi, de la documentation technique, et du suivi qualité, écriture des logiciels nécessaires pour le produit...*
- *Coûts immatériels au niveau de l'atelier de production : cercles de qualité, programmation des machines-outils à commandes numériques, choix techniques réalisés par les opérateurs (définition des paramètres d'usinage, choix des matériaux,...)*
- *Coûts de commercialisation : publicité, conseil au client, formation des clients, négociation du prix et des clauses du contrat...*
- *Coûts liés à la logistique : gestion et organisation du transport et du stockage, comptabilité, facturation, ...*
- *Coûts du services après vente: maintenance, contentieux,...*
- *Coût des informations que l'on achète : brevets, licences, accès à des banques de données,..."*

La part de la documentation, du point de vue du contenu comme des méthodes et des systèmes permettant de l'exploiter, est importante puisqu'elle intervient au moins à trois niveaux, voire quatre :

- Les études préliminaires de marché
- La conception
- Les choix techniques (choix de matière par exemple)
- Si on fait prendre en compte la circulation, la mémorisation et l'archivage de l'ensemble des documents produits dans une organisation (fonction que s'attribue la GED : "Gestion Electronique des Documents"), la part de la documentation sous-tend une grande partie de

ce qui constitue la dimension immatérielle d'un produit car toutes les fonctions immatérielles sont confrontées à la gestion des documents. Il faut ajouter le rôle d'Internet dans ce domaine qui, grâce aux possibilités croisées de la messagerie, du FTP (File Transfer Protocol) et du plus grand système documentaire, le WEB, rend possible la diffusion et l'échange des documents sur un réseau ouvert comme sur un réseau fermé (Intranet et Extranet). Cette dimension apportée par Internet à la gestion électronique des documents est susceptible de réduire les coûts informationnels d'une manière assez considérable. C'est la raison d'être de projets tels le programme américain de l'Automotive Exchange Network, consistant dans un gigantesque Intranet reliant tous les acteurs de la filière automobile afin de réduire les délais liés à la circulation des documents.

Nous disposons donc aujourd'hui de moyens de diffusion et d'échange de documents considérables grâce à des réseaux de tout type : réseau local, réseau planétaire (Internet et ses versions Extranet/Intranet) et les protocoles associés. Les questions relatives à la normalisation des documents sur Internet en particulier, conduisant à une indexation plus fine par les moteurs de recherche, restent en cours d'élaboration.

Nous disposons également de méthodes de traitement de la documentation intéressantes grâce à des puissances croissantes de calcul permettant d'obtenir des résultats sur de grandes quantités de documents et relativement peu coûteuses. Les logiciels sont peu onéreux, les calculs peuvent se faire sur de puissants micro-ordinateurs donc à des prix accessibles. Sur le plan du hardware, le coût des serveurs est en baisse croissante et permet de diffuser de l'information à des prix raisonnables. Il reste, de la part des industriels en particulier, la motivation pour investir dans l'information. Les besoins d'information, notamment dans le domaine de la plasturgie pour lequel nous avons réalisé ces travaux, évoluent. Mais demeure une incertitude quant au gain obtenu en contrepartie de l'investissement en matière d'information. Cette incertitude est liée à l'un des enjeux majeurs d'un programme de développement de l'intelligence économique : une politique de diffusion des contenus qui soit à la hauteur de ses ambitions.

Les pouvoirs publics s'attèlent à cette tâche mais ils restent en retrait par rapport à la politique américaine de diffusion de l'information. Nous avons vu, durant les trois dernières années de ce travail, l'information-brevet (américaine) apparaître en accès libre sur Internet (l'information-brevet étant pourtant l'une des plus onéreuses sur le marché),

Medline (base de données de référence en médecine) devenir gratuite dans sa version complète sur Internet.

Il nous semble également que les organisations professionnelles ont un rôle considérable à jouer dans le domaine de la diffusion des contenus et des compétences en matière de traitement de l'information, rôle qu'elles n'ont pas véritablement joué. Le simple accès à des données telles que "la consommation de telle matière plastique dans l'industrie automobile en Europe" ou "le chiffre d'affaire de l'industrie du moule en Europe" n'est pas accessible directement à l'industriel plasturgiste. Or l'accès à ces données serait la prémisse d'une gestion stratégique de l'information dont l'étape suivante serait la mise à disposition pour les industriels de véritables systèmes de renseignement (accès aux données et traitement des documents) permettant un réel développement de l'intelligence économique.

## **RESULTATS DETAILLES**

## **Résultats détaillés**

Les résultats sont issus *du traitement des résumés des articles*, seuls textes disponibles en matière d'information scientifique et technique. Pour une présentation des résultats plus simple et moins longue, nous avons parfois seulement cité les titres des documents *sans* le résumé. Quand nous l'avons fait, les titres étaient eux-mêmes suffisamment explicites pour illustrer la cohérence de la métaclé de telle sorte que la simple lecture des titres des documents était suffisante à le montrer sans qu'il soit besoin d'éditer l'ensemble des résumés.

**RESULTATS N°2**  
**POLYMERES A CRISTAUX LIQUIDES**

**Le ciblage des données marché**

*Métaclé et texte des résumés*

RAPPEL DE LA METACLE

IL Y A 18 MOTS EXPLICATIFS :POUR LE FACTEUR 1 LISTE DES MOTS A COORDONNEE NEGATIVE

SEUILS POUR LA CONTRIBUTION 2\*MOY = 6.76 3\*MOY = 10.14

LIGNE	FACTEUR	QUALITE	CONTRI		
40	-0.67	99.09	14.72	3	VECTRA
41	-0.77	68.32	9.62	19	GIVEN
42	-0.87	92.47	13.61	14	APPLICATIONS
42	-0.88	47.02	7.37	46	RESISTANCE
43	-0.97	65.14	10.57	36	KYDAR
44	-1.04	170.41	21.46	11	CELANESE
44	-1.02	164.54	20.49	12	DETAILS
44	-0.97	69.96	11.80	32	PLASTICS
45	-1.06	127.81	15.79	24	HOECHST
45	-1.11	117.30	15.61	30	BRIEF
45	-1.10	74.77	10.84	47	COMPANY
45	-1.06	49.87	8.62	53	GRADES
45	-1.05	56.14	8.55	54	PRODUCTS
45	-1.09	52.39	7.53	68	MARKET
46	-1.15	48.56	7.49	75	DARTCO
47	-1.24	65.18	9.44	71	NOTED
47	-1.22	59.27	8.06	79	PLANT
48	-1.31	48.03	6.98	102	JAPAN

Liste des 9 documents du FACTEUR 1 Coordonnées Négatives au niveau de qualité [ 67 , 250 [

Les mots de la métaclé sont soulignés dans les textes.

**Résumé Document 00001044 ==> Ligne 41 Fac = -0.78 Qualité = 73.73 Contribution = 2.47**

**TI - VECTRA LCP**

**SO - Eur.Plast.News; 13, No.6, June 1986, p.46**

**AB - Excellent chemical resistance is the reason why a change from ceramic saddles to a new tower packing system injection moulded from Celanese's Vectra liquid crystal polymer is now saving more than 110, 000 US dollars/yr. at a formic acid plant in the USA. Produced by Norton Chemical Process Products, the tower packings are said to be highly impermeable and resistant to high temperature and corrosion; brief details are given.**

**Résumé Document 00000924 ==> Ligne 43 Fac = -0.95 Qualité = 69.40 Contribution = 1.82**

**TI - VECTRA LCP MEETS CD REQUIREMENTS**

**SO - Plast.Rubb.Wkly.; No.1285, 13th May 1989, p.12**

**AB - Hoechst Celanese Plastics' high-modulus Vectra liquid crystal polymer has been selected by Philips Industrie for use as a component in its CD players. Lens holder mouldings are made from a glass-filled grade of Vectra with a specific modulus of between 15 and 17.**

**Résumé Document 00000532 ==> Ligne 44 Fac = -1.02 Qualité = 97.77 Contribution = 7.75**

**TI - LIQUID CRYSTAL POLYMER**

**SO - Mod.Plast.Int.; 25, No.1, Jan.1995, p.48/50**

**AB - Hoechst Celanese has reaffirmed its commitment to the liquid crystal polymer segment by unveiling new capacity for its Vectra LCP. To meet the demand from fast-growing Asian markets, a 3,000 t/y plant will be built in Japan by Polyplastics, a Hoechst joint venture with Daicel Chemical Industries. Asian demand is expected to grow to approximately 910 tonnes by the end of the decade. DuPont has started construction of a commercial-scale facility at an existing site in Chattanooga and has named its melt-processable products Zenite. Amoco Polymers is targeting electronics applications with its newest grade, Xydar G-930. The glass-reinforced injection moulding resin is said to fill thin walls over long flow lengths with little or no flash. Despite their cost, LCPs are enjoying 25% annual growth worldwide and are forecasted to maintain that level through the decade. The global market for LCPs is about 4,500 tonnes.**

**Résumé Document 00001012 ==> Ligne 44 Fac = -1.02 Qualité = 80.65 Contribution = 2.33**

**TI - LIQUID CRYSTAL POLYMER RATED FOR ELECTRONICS**

**SO - Plast.S.Africa; 16, No.6, April 1987, p.49**

**AB - Vectra thermotropic liquid crystal polymer from Celanese is a new material particularly suitable for surface mount applications in the electronics industry. Brief details are given of Vectra's properties.**

**Résumé Document 00001052 ==> Ligne 44 Fac = -0.99 Qualité = 71.37 Contribution = 4.25**

**TI - MORE EUROPE INVOLVEMENT PROMISED BY CELANESE**

**SO - Plast.Rubb.Wkly.; No.1134, 19th April 1986, p.9**

**AB - Engineering thermoplastics major Celanese is committed to consolidating and expanding its European Investment, and Whitfield & Son, the Telford-based compounder purchased in the New Year from English China Clays, is to play a major role in this expansion programme. Also important is the Business Concepts Group running in Watford whose aim is to develop new engineering resins and applications for them. Brief details are given of new products introduced by Celanese including Durel**

polycarbonate, Duraloy polyester and polyacetal blends, Vectra, a liquid crystal polymer and Kematal electro-plating grades.

**Résumé Document 00001061 ==> Ligne 45 Fac = -1.11 Qualité = 71.60 Contribution = 2.76**

**TI - LIQUID-CRYSTAL POLYMERS: NOW THEY ARE MELT-PROCESSABLE**  
**SO - Mod.Plast.Int.; 15, No.9, Sept.1985, p.66-8**

**AB - Details are given of the first melt-processable liquid-crystal polymer to come onto the market, Dartco's Xydar self-reinforcing aromatic copolyester, and of different grades of an aromatic polyester which Celanese expect to introduce later in 1985. Processing and properties of liquid-crystal polymers are described in detail. Xydar is produced from terephthalic and p-hydroxybenzoic acids and biphenol. Mouldable in regular tooling and on standard equipment, the polymers offer spectacular properties without recourse to abrasive reinforcements.**

**Résumé Document 00000962 ==> Ligne 46 Fac = -1.20 Qualité = 67.56 Contribution = 2.95**

**TI - NEW LIQUID CRYSTAL RESINS FOR E/E USES**  
**SO - Plast.Technol.; 34, No.10, Sept.1988, p.87/9**

**AB - At the NPE exhibition in Chicago, Amoco Performance Products is reported to have not only featured its newly-acquired line of Xydar liquid crystal polymers, but introduced eight new liquid crystal polymer grades said to be particularly well suited to electrical/electronic applications. Details of properties are given.**

**Résumé Document 00000964 ==> Ligne 47 Fac = -1.27 Qualité = 88.72 Contribution = 4.73**

**TI - LIQUID-CRYSTAL POLYMER DEVELOPERS OFFER SYSTEMS WITH EASIER PROCESSING CHARACTERISTICS**  
**SO - Adv.Mat.(Newsletter); 10, No.10, 27th June 1988, p.2-3**

**AB - Du Pont is reported to have entered the liquid crystal polymer market with five grades: Amoco, which recently bought Dartco's Xydar range of materials has introduced several more easily processable versions; and Hoechst Celanese has continued to expand its Vectra line of liquid crystal polymers. Brief details of each product are included.**

**Résumé Document 00000969 ==> Ligne 47 Fac = -1.22 Qualité = 72.33 Contribution = 3.01**

**TI - VECTRA LIQUID-CRYSTAL POLYMER TO BE PRODUCED HERE**  
**SO - Jap.Chem.Week; 29, No.1474, 14th July 1988, p.3**

**AB - Properties and applications for 'Vectra' liquid-crystal polymers are briefly given in this short article, which announces that the Vectra polymers are to be produced in Japan by Hoechst Celanese's subsidiary Polyplastics Co.**

## RESULTATS N°4

### Exemple de typologie thématique :

*Les polymères à cristaux liquides (LCP)* · 548 documents

Les neufs thèmes principaux de l'une des analyses

A chaque classe est associée une dizaine de documents, quelque soit la façon de réaliser l'analyse factorielle. Nous donnons ci-après l'une de ces analyses avec, pour le premier facteur et quelques autres la liste des textes, et pour la suite seulement la liste des titres quand elle suffit en général à faire ressortir le thème de la classe. Toutefois nous rappelons que l'analyse factorielle a été faite sur les résumés des articles.

#### 1. Les données commerciales

Résumé N° 00001044

TI - VECTRA LCP

SO - Eur.Plast.News; 13, No 6, June 1986, p 46

DT - Journal

LA - English

AB - Excellent chemical resistance is the reason why a change from ceramic saddles to a new tower packing system injection moulded from Celanese's Vectra liquid crystal polymer is now saving more than 110,000 US dollars/yr at a formic acid plant in the USA. Produced by Norton Chemical Process Products, the tower packings are said to be highly impermeable and resistant to high temperature and corrosion; brief details are given.

CO - CELANESE CORP.; NORTON CHEMICAL PROCESS PRODUCTS

PY - 1986

Résumé N° 00000924

TI - **VECTRA LCP MEETS CD REQUIREMENTS**

SO - Plast Rubb Wkly, No 1285, 13th May 1989, p 12

DT - Journal

LA - English

AB - Hoechst Celanese Plastics' high-modulus Vectra liquid crystal polymer has been selected by Philips Industrie for use as a component in its CD players. Lens holder mouldings are made from a glass-filled grade of Vectra with a specific modulus of between 15 and 17.

CO - HOECHST CELANESE PLASTICS; PHILIPS INDUSTRIE NV

PY - 1989

Résumé N0 00000532

**TI - LIQUID CRYSTAL POLYMER**

AU - Grande J A ,

SO - Mod.Plast Int , 25, No 1, Jan 1995, p 48/50

DT - Journal

LA - English

AB - Hoechst Celanese has reaffirmed its commitment to the liquid crystal polymer segment by unveiling new capacity for its Vectra LCP. To meet the demand from fast-growing Asian markets, a 3,000 t/y plant will be built in Japan by Polyplastics, a Hoechst joint venture with Daicel Chemical Industries. Asian demand is expected to grow to approximately 910 tonnes by the end of the decade. DuPont has started construction of a commercial-scale facility at an existing site in Chattanooga and has named its melt-processable products Zenite. Amoco Polymers is targeting electronics applications with its newest grade, Xydar G-930. The glass-reinforced injection moulding resin is said to fill thin walls over long flow lengths with little or no flash. Despite their cost, LCPs are enjoying 25% annual growth worldwide and are forecasted to maintain that level through the decade. The global market for LCPs is about 4,500 tonnes.

PY - 1995

Résumé N0 00001012

**TI - LIQUID CRYSTAL POLYMER RATED FOR ELECTRONICS**

SO - Plast S Africa; 16, No 6, April 1987, p 49

DT - Journal

LA - English

AB - Vectra thermotropic liquid crystal polymer from Celanese is a new material particularly suitable for surface mount applications in the electronics industry. Brief details are given of Vectra's properties.

CO - CELANESE CORP

PY - 1987

Résumé N0 00001052

**TI - MORE EUROPE INVOLVEMENT PROMISED BY CELANESE**

SO - Plast Rubb Wkly , No 1134, 19th April 1986, p 9

DT - Journal

LA - English

AB - Engineering thermoplastics major Celanese is committed to consolidating and expanding its European Investment, and Whitfield & Son, the Telford-based compounder purchased in the New Year from English China Clays, is to play a major role in this expansion programme. Also important is the Business Concepts Group running in Watford whose aim is to develop new engineering resins and applications for them. Brief details are given of new products introduced by Celanese including Durel polycarbonate, Duraloy polyester and polyacetal blends, Vectra, a liquid crystal polymer and Kematal electro-plating grades.

CO - CELANESE CORP ; POLYMER COMPOSITES INC

PY - 1986

Résumé N0 00001061

**TI - LIQUID-CRYSTAL POLYMERS: NOW THEY ARE MELT-PROCESSABLE**

AU - Wood A S ,

SO - Mod.Plast Int ; 15, No 9, Sept 1985, p 66-8

DT - Journal

LA - English

AB - Details are given of the first melt-processable liquid-crystal polymer to come onto the market, Dartco's Xydar self-reinforcing aromatic copolyester, and of different grades of an aromatic polyester which Celanese expect to introduce later in 1985. Processing and properties of liquid-crystal polymers are described in detail. Xydar is produced from terephthalic and p-hydroxybenzoic acids and biphenol. Mouldable in regular tooling and on standard equipment, the polymers offer spectacular properties without recourse to abrasive reinforcements.

CO - CELANESE CORP., DARTCO MANUFACTURING INC  
PY - 1985

Résumé N0 00000962

**TI - NEW LIQUID CRYSTAL RESINS FOR E/E USES**

SO - Plast Technol , 34, No 10, Sept 1988, p 87/9

DT - Journal

LA - English

AB - At the NPE exhibition in Chicago, Amoco Performance Products is reported to have not only featured its newly-acquired line of Xydar liquid crystal polymers, but introduced eight new liquid crystal polymer grades said to be particularly well suited to electrical/electronic applications. Details of properties are given.

CO - AMOCO PERFORMANCE PRODUCTS CO  
PY - 1988

Résumé N0 00000964

**TI - LIQUID-CRYSTAL POLYMER DEVELOPERS OFFER SYSTEMS WITH EASIER PROCESSING CHARACTERISTICS**

SO - Adv Mat (Newsletter); 10, No 10, 27th June 1988, p 2-3

DT - Journal

LA - English

AB - Du Pont is reported to have entered the liquid crystal polymer market with five grades. Amoco, which recently bought Dartco's Xydar range of materials has introduced several more easily processable versions; and Hoechst Celanese has continued to expand its Vectra line of liquid crystal polymers. Brief details of each product are included.

CO - AMOCO CHEMICALS CORP , DU PONT DE NEMOURS E I , è CO INC ,  
HOECHST CELANESE CORP  
PY - 1988

Résumé N0 00000969

**TI - VECTRA LIQUID-CRYSTAL POLYMER TO BE PRODUCED HERE**

SO - Jap Chem Week; 29, No 1474, 14th July 1988, p 3

DT - Journal

LA - English

AB - Properties and applications for 'Vectra' liquid-crystal polymers are briefly given in this short article, which announces that the Vectra polymers are to be produced in Japan by Hoechst Celanese's subsidiary Polyplastics Co.

CO - HOECHST CELANESE CORP ; POLYPLASTICS CO  
PY - 1988

2. Les données matériau · études des propriétés du matériau en fonction de la phase dans lequel il se trouve (smectique, nématique , cholesterique), ces phases se caractérisant par l'arrangement des molécules entre elles Les polymères à cristaux liquides sont en effet des liquides qui s'arrangent comme des solides · ils ont une structure ordonnée Leurs propriétés dépendent de cette structure

FACTEUR 1 Coordonnées Positives au niveau de qualité [ 67 , 250 ]

Résumé Document 00000734 ==> Ligne 6 Fac = 2 08 Qualité = 88 77 Contribution = 5 86

**TI - KINETICS OF ORIENTATION OF A SIDE-CHAIN LIQUID-CRYSTAL POLYMER AS MEASURED BY A HIGH-TEMPERATURE SQUID-MAGNETOMETER**

SO - J Polym Sci.Polym Phys., 30, No 11, Oct 1992, p 1199-206

Résumé Document 00000788 ==> Ligne 7 Fac = 1 97 Qualité = 67 39 Contribution = 7 03

**TI - DYNAMIC LIGHT SCATTERING FROM A SIDE-CHAIN LIQUID CRYSTAL POLYMER IN A NEMATIC SOLVENT**

SO - Miami Beach, Fl , Fall 1989, p.11-5 012

Résumé Document 00000589 ==> Ligne 8 Fac = 1 88 Qualité = 87 19 Contribution = 17 60

**TI - RHEOLOGICAL CHARACTERISATION OF DIRECTOR DYNAMICS IN A NEMATIC MONODOMAIN CONTAINING MESOGENIC POLYMERS OF DIFFERING ARCHITECTURES**

SO - Macromolecules; 27, No 2, 17th Jan 1994, p 337-47

Résumé Document 00000582 ==> Ligne 10 Fac = 1 73 Qualité = 79 63 Contribution = 10 80

**TI - TEMPERATURE AND MOLECULAR WEIGHT DEPENDENCE OF THE VISCOMETRIC PROPERTIES OF MAIN-CHAIN LIQUID CRYSTAL POLYMERS IN NEMATIC SOLVENTS**

SO - Macromolecules; 27, No.17, 15th Aug 1994, p 4691-8

Résumé Document 00000661 ==> Ligne 11 Fac = 1 69 Qualité = 82 29 Contribution = 14 88

**TI - EFFECTS OF NON-MESOGENIC COMONOMERS ON THE PHASE BEHAVIOUR, THE STRUCTURE AND THE MISCIBILITY OF SIDE CHAIN LIQUID CRYSTALLINE POLYMERS**

SO - Makromol Chem , 194, No 8, Aug 1993, p 2349-59

Résumé Document 00000687 ==> Ligne 11 Fac = 1 62 Qualité = 77 10 Contribution = 12 49

**TI - INVESTIGATION OF THE SMECTIC-ISOTROPIC TRANSITION IN A SIDE-CHAIN LIQUID CRYSTAL POLYMER BY SYNCHROTRON RADIATION X-RAY DIFFRACTION**

SO - J Polym Sci.Polym.Phys , 31, No 7, June 1993, p 773-7

Résumé Document 00000691 ==> Ligne 11 Fac = 1 68 Qualité = 67 65 Contribution = 7 05

**TI - CONSTITUTIVE EQUATION FOR LIQUID CRYSTALLINE POLYMER SOLUTIONS**

SO - J.Rheol , 37, No 3, May/June 1993, p.413-41

Résumé Document 00000542 ==> Ligne 12 Fac = 1 60 Qualité = 95 06 Contribution = 4 04

**TI - INDUCED CHOLESTERIC LIQUID CRYSTAL POLYMER AS A NEW MEDIUM FOR OPTICAL DATA STORAGE**

SO - Molecular Crystals è Liquid Crystals, Vol 251, 1994, p 311-6

Résumé Document 00000705 ==> Ligne 12 Fac = 1 55 Qualité = 82 14 Contribution = 8 73

**TI - PROTON NMR SPECTROSCOPIC INVESTIGATION OF THE DYNAMICS OF DIRECTOR ROTATION IN A COMB-SHAPED LIQUID CRYSTAL POLYMER**

SO - Polymer Science, 35, No 2, Feb 1993, p 183-7

Résumé Document 00000615 ==> Ligne 13 Fac = 1 52 Qualité = 77 05 Contribution = 10 41

**TI - MOLECULAR WEIGHT DEPENDENT BEHAVIOUR OF THE TWIST DISTORTION IN A NEMATIC MONODOMAIN CONTAINING A MAIN-CHAIN LIQUID CRYSTAL POLYMER**

SO - Macromolecules; 27, No 7, 28th March 1994, p 1943-8

Résumé Document 00000791 ==> Ligne 13 Fac = 1 49 Qualité = 72 54 Contribution = 9 47

**TI - MAIN-CHAIN SIDE-CHAIN LIQUID CRYSTAL POLYMER BLENDS FOR IMPROVED PHYSICAL PROPERTIES**

SO - Makromol Chem ,Macromol Symp.; Vol 51, Oct 1991, p 41-60

Résumé Document 00000522 ==> Ligne 14 Fac = 1 40 Qualité = 83 40 Contribution = 14 99

**TI - PHASE SEPARATION IN LIQUID CRYSTAL/UV CURABLE MONOMER SYSTEMS DURING POLYMERISATION**

SO - Kobunshi Ronbun ; 51, No 5, May 1994, p 303-7

Résumé Document 00000657 ==> Ligne 14 Fac = 1 38 Qualité = 108 21 Contribution = 7 32

**TI - TRANSESTERIFICATION IN NEMATIC POLYMERS**

SO - Macromolecules, 26, No.17, 16th Aug 1993, p 4499-505

Résumé Document 00001060 ==> Ligne 14 Fac = 1 45 Qualité = 81 47 Contribution = 2 86

**TI - STUDY OF THE STATIC AND DYNAMIC PROPERTIES OF SIDE-CHAIN LIQUID CRYSTALLINE POLYMERS IN LOW MOLAR MASS MESOGENS**

SO - Polymer; 26, No 9, Aug.1985, p 1319-24

Résumé Document 00000541 ==> Ligne 15 Fac = 1 33 Qualité = 100 41 Contribution = 5 63

**TI - CHARACTERISATION OF ORDERING IN A MAIN CHAIN LIQUID CRYSTAL POLYMER BY SMALL ANGLE NEUTRON SCATTERING AND NMR**

SO - Macromolecular Reports; A31, Nos 6 è 7, Oct 1994, p 1101-108

Résumé Document 00000672 ==> Ligne 16 Fac = 1.25 Qualité = 81.15 Contribution = 6 03

**TI - CHAIN CONFORMATIONS OF LIQUID CRYSTALLINE POLYMERS**

SO - Polymer; 34, No 14, July 1993, p 3112-4

Résumé Document 00000694 ==> Ligne 16 Fac = 1.27 Qualité = 83 96 Contribution = 2 91

**TI - PHOTO-FRIES REARRANGEMENT IN A SIDE-CHAIN LIQUID CRYSTALLINE POLYMER**

SO - Polymer, 34, No 7, 1993, p 1347-53

3 Propriétés, dans la transformation, des mélanges ("blends") à base de polymère à cristaux liquides.

FACTEUR 2 Coordonnées Négatives au niveau de qualité [ 67 , 250 [

Résumé Document 00000629 ==> Ligne 35 Fac = -0.80 Qualité = 83 90 Contribution = 3 61

**TI - DEVELOPMENT OF LIQUID CRYSTALLINE POLYESTER COMPOSITE AS RECYCLABLE HIGH-STRENGTH MATERIAL**

SO - Kobunshi Ronbun , 50, No 11, 1993, p 855-62

Résumé Document 00000807 ==> Ligne 35 Fac = -0 79 Qualité = 88 71 Contribution = 3 48

**TI - RHEOLOGY, MORPHOLOGY, AND MECHANICAL CHARACTERISTICS OF POLY(ETHERETHER KETONE) - LIQUID CRYSTAL POLYMER BLENDS**

SO - Polym Engng Sci ; 31, No.13, Mid-July 1991, p 971-80

Résumé Document 00000689 ==> Ligne 36 Fac = -0 92 Qualité = 117 27 Contribution = 4 92

**TI - IMPROVED PROCESSABILITY BY BLENDING THERMOPLASTICS WITH LIQUID CRYSTAL POLYMERS**

SO - Macplas Int , Oct 1992, p 89-91

Résumé Document 00000700 ==> Ligne 36 Fac = -0 87 Qualité = 73 17 Contribution = 3 39

**TI - EFFECTS OF THE ADDITION OF A LIQUID CRYSTALLINE COPOLYESTER TO POLYSTYRENES ON BLENDING TORQUE AND MECHANICAL PROPERTIES OF BLENDS**

SO - J Appl Polym Sci ; 48, No.3, 15th April 1993, p.383-91

Résumé Document 00000826 ==> Ligne 36 Fac = -0 92 Qualité = 70 12 Contribution = 4 71

**TI - STRUCTURE AND PROPERTIES OF MOULDED POLYBLENDS CONTAINING LIQUID CRYSTALLINE POLYMERS**

SO - Polym Engng Sci., 31, No 6, March 1991, p 459-66

Résumé Document 00000722 ==> Ligne 37 Fac = -1 01 Qualité = 112 60 Contribution = 5 45

**TI - PROPERTIES OF BLENDS OF A THERMOTROPIC LIQUID CRYSTALLINE POLYMER WITH A FLEXIBLE POLYMER (VECTRA/PETP)**

SO - J Rheol , 36, No 8, Nov.1992, p 1477-98

Résumé Document 00000806 ==> Ligne 37 Fac = -0.96 Qualité = 75.32 Contribution = 4 64

**TI - PROCESSING AND PROPERTIES OF POLYCARBONATE/LIQUID CRYSTAL POLYMER BLENDS**

SO - Int Polym Processing, 6, No 3, Sept 1991, p 247-52

Résumé Document 00000816 ==> Ligne 37 Fac = -0 97 Qualité = 89 49 Contribution = 3 70

**TI - BLENDS OF A LIQUID CRYSTALLINE POLYMER WITH PEEK**

SO - Dallas, Tx., 7th-11th May 1990, p.489-92. 012

Résumé Document 00000745 ==> Ligne 38 Fac = -1 09 Qualité = 119 57 Contribution = 5 64  
**TI - EXTRUDED BLENDS OF A THERMOTROPIC LIQUID CRYSTALLINE POLYMER WITH POLYETHYLENE TEREPHTHALATE, POLYPROPYLENE AND POLYPHENYLENE SULPHIDE**  
SO - J Appl.Polym Sci , 44, No 12, 25th April 1992, p 2185-95

Résumé Document 00000748 ==> Ligne 38 Fac = -1 12 Qualité = 139.04 Contribution = 5 93  
**TI - PROCESSING OF LIQUID CRYSTAL/ENGINEERING POLYMER BLENDS**  
SO - Macplas; 17, No.137, April 1992, p 95-6

Résumé Document 00000778 ==> Ligne 38 Fac = -1.05 Qualité = 75 17 Contribution = 4 63  
**TI - BLENDS OF A LIQUID CRYSTALLINE POLYMER WITH POLYETHER ETHER KETONE**  
SO - Polym Engng Sci , 32, No 2, Jan 1992, p 85-93

Résumé Document 00000809 ==> Ligne 38 Fac = -1 05 Qualité = 73 29 Contribution = 4 60  
**TI - STUDIES ON A DIPHENYLETHER MODIFIED POLY(PHENYL-1,4-PHENYLENE TEREPHTHALATE) LIQUID CRYSTALLINE POLYMER AND ITS BLENDS WITH POLYCARBONATE AND POLYSULPHONE**  
SO - Polym Engng.Sci ; 31, No 13, Mid-July 1991, p 954-62

Résumé Document 00000529 ==> Ligne 39 Fac = -1.15 Qualité = 92 19 Contribution = 5 12  
**TI - INFLUENCE OF PROCESSING HISTORY ON THE PROPERTIES OF A THERMOTROPIC COPOLYESTER/POLYCARBONATE BLEND**  
SO - Polym Engng Sci., 35, No 1, Mid-Jan 1995, p 52-63

Résumé Document 00000566 ==> Ligne 39 Fac = -1 16 Qualité = 94 40 Contribution = 8 18  
**TI - THERMAL AND MECHANICAL PROPERTIES OF INJECTION MOULDED BLENDS OF A LIQUID CRYSTALLINE POLYMER AND POLY(BUTYLENE TEREPHTHALATE)**  
SO - Polym Engng Sci , 34, No 17, Mid-Sept.1994, p 1346-53

Résumé Document 00000583 ==> Ligne 41 Fac = -1 41 Qualité = 73 23 Contribution = 7 20  
**TI - PROCESSABILITY AND SHEAR STABILITY OF THE BLENDS OF THREE LIQUID CRYSTALLINE POLYMERS WITH POLYAMIDEIMIDE**  
SO - New Orleans, La , 9th-13th May 1993, Vol III, p 2676-85 012

Résumé Document 00000636 ==> Ligne 41 Fac = -1 36 Qualité = 79 49 Contribution = 14 94  
**TI - MECHANISM OF SKIN-CORE MORPHOLOGY FORMATION IN EXTRUDATES OF POLYCARBONATE/LIQUID CRYSTALLINE POLYMER BLENDS**  
SO - Polym Engng Sci., 33, No 23, Mid-Dec 1993, p.1548-58

Résumé Document 00000777 ==> Ligne 41 Fac = -1.40 Qualité = 86 61 Contribution = 18 55  
**TI - SHEAR AND ELONGATIONAL VISCOSITIES OF A THERMOTROPIC LIQUID CRYSTALLINE POLYMER AND ITS BLEND WITH POLYCARBONATE**  
SO - Polym Engng.Sci ; 32, No 1, Mid-Jan 1992, p 14-9

4. LCP et fibres en particulier fibres de carbone.

**FACTEUR 3 Coordonnees Negatives au niveau de qualité : [ 67 , 250 ]**

**Résumé Document 00000991 ==> Ligne 37 Fac = -1.26 Qualité = 71.91 Contribution = 5.61**

**TI - FIBROUS REINFORCEMENTS: THE CHOICES ARE GETTING TOUGHER**  
SO - Mod Plast Int ; 17, No.9, Sept 1987, p.56-8

**Résumé Document 00000651 ==> Ligne 39 Fac = -1.47 Qualité = 104.44 Contribution = 9.08**

**TI - DEVELOPMENT AND EVALUATION OF SURFACE TREATMENTS TO ENHANCE THE FIBRE-MATRIX ADHESION IN PAN-BASED CARBON FIBRE/LIQUID CRYSTAL POLYMER COMPOSITES. II. ELECTROCHEMICAL TREATMENTS**

SO - Polym Composites, 14, No 4, Aug 1993, p 301-7

**Résumé Document 00000650 ==> Ligne 40 Fac = -1.60 Qualité = 77.77 Contribution = 10.86**

**TI - GRAFTING OF A FUNCTIONALISED SIDE-CHAIN LIQUID CRYSTAL POLYMER ON CARBON FIBRE SURFACES: NOVEL COUPLING AGENTS FOR FIBRE/POLYMER MATRIX COMPOSITES**

SO - Chem of Mat , 5, No 9, Sept 1993, p 1287-92

**Résumé Document 00000881 ==> Ligne 40 Fac = -1.59 Qualité = 91.67 Contribution = 15.14**

**TI - LIQUID CRYSTAL POLYMER-CARBON FIBRE COMPOSITES. MOLECULAR ORIENTATION**

SO - Polym Engng Sci., 30, No 4, Feb 1990, p 228-34

**Résumé Document 00000652 ==> Ligne 41 Fac = -1.85 Qualité = 142.01 Contribution = 15.59**

**TI - DEVELOPMENT AND EVALUATION OF SURFACE TREATMENTS TO ENHANCE THE FIBRE-MATRIX ADHESION IN PAN-BASED CARBON FIBRE/LIQUID CRYSTAL POLYMER COMPOSITES. I. COUPLING AGENT AND AMINE SURFACE TREATMENTS**

SO - Polym Composites, 14, No 4, Aug 1993, p 292-300

**Résumé Document 00000747 ==> Ligne 41 Fac = -1.74 Qualité = 71.98 Contribution = 11.69**

**TI - MEASUREMENT AND ANALYSIS OF THE FRACTURE OF A LIQUID CRYSTAL POLYMER**

SO - J Mat Sci ; 27, No 14, 15th July 1992, p 3969-76

**Résumé Document 00001037 ==> Ligne 41 Fac = -1.73 Qualité = 123.26 Contribution = 17.97**

**TI - INFLUENCE OF SHORT-FIBRE REINFORCEMENT ON THE FRACTURE BEHAVIOUR OF A BULK LIQUID CRYSTAL POLYMER**

SO - J Mat Sci , 24, No 8, Aug 1986, p 2889-900

**Résumé Document 00001050 ==> Ligne 41 Fac = -1.80 Qualité = 105.81 Contribution = 24.93**

**TI - INDUCED ORIENTATIONAL BEHAVIOUR OF LIQUID CRYSTAL POLYMER BY CARBON FIBRES**

SO - Polym.Composites, 6, No 3, July 1985, p 181-4

**Résumé Document 00000581 ==> Ligne 42 Fac = -1.98 Qualité = 109.28 Contribution = 4.15**

**TI - COMPOSITE PRINTED CIRCUIT BOARD SUBSTRATE AND PROCESS FOR ITS MANUFACTURE**

**Résumé Document 00000868 ==> Ligne 42 Fac = -1.90 Qualité = 99.85 Contribution = 19.09**

**TI - GRAFTING OF A LIQUID CRYSTAL POLYMER ON CARBON FIBRES**

SO - Polym Engng Sci ; 30, No 10, May 1990, p 603-8

**Résumé Document 00000545 ==> Ligne 43 Fac = -2.09 Qualité = 140.74 Contribution = 58.49**

**TI - COUPLING CARBON FIBRES TO EPOXY MATRICES WITH GRAFTED SIDE CHAIN LIQUID CRYSTAL POLYMERS**

SO - Chem of Mat , 6, No.10, Oct.1994, p 1880-3

**Résumé Document 00000587 ==> Ligne 43 Fac = -2.13 Qualité = 108.13 Contribution = 50.87**

**TI - MORPHOLOGY OF A THERMOTROPIC RANDOM TERPOLYMER LIQUID CRYSTAL POLYMER CRYSTALLISED IN THE BULK. COMPRESSION MOULDINGS, EXTRUDATES AND FIBRES**

SO - Polymer, 35, No 9, 1994, p 1869-79

5 Le problème des lignes de soudure dans la fabrication des pièces en LCP

**FACTEUR 4 Coordonnées Positives au niveau de qualité : [ 67 , 250 ]**

**Résumé Document 00000505 ==> Ligne 8 Fac = 2.66 Qualité = 96.34 Contribution = 54.27**

**TI - ENHANCEMENT OF INTERNAL WELD LINE STRENGTH IN THERMOTROPIC LIQUID CRYSTAL POLYMER MOULDINGS**

SO - Plast.Rubb.Comp.Process.Appln.; 23, No.3, 1995, p.139-50

**Résumé Document 00000595 ==> Ligne 15 Fac = 1.71 Qualité = 86.59 Contribution = 15.04**

**TI - INFLUENCE OF GLASS FIBRES ON THE SPIN WELDING OF THERMOTROPIC LIQUID CRYSTALLINE POLYMER**

SO - New Orleans, La., 9th-13th May 1993, Vol.II, p.2082-8. 012

**Résumé Document 00000621 ==> Ligne 16 Fac = 1.47 Qualité = 104.74 Contribution = 14.19**

**TI - IMPROVING WELD LINE STRENGTHS USING A MULTI-LIVE FEED INJECTION MOULDING PROCESS**

SO - New Orleans, La., 9th-13th May 1993, Vol.I, p.521-9. 012

**Résumé Document 00001030 ==> Ligne 18 Fac = 1.20 Qualité = 99.27 Contribution = 8.40**

**TI - POLYMER PROCESSING MACHINERY. THE APPLICATION OF OSCILLATING PACKING PRESSURES IN THICK-SECTION INJECTION MOULDING**

**SO - (Bristol), c.1985, pp.9. 12ins. 4/3/86. 831-9(12)1**

**Résumé Document 00000596 ==> Ligne 19 Fac = 1.13 Qualité = 80.45 Contribution = 6.05**

**TI - IMPROVING WELD LINE STRENGTHS USING A MULTI-LIVE FEED INJECTION MOULDING PROCESS**

**SO - Rochester, NY, 15th-16th Sept.1993, paper W, pp.15. 012**

**Résumé Document 00000911 ==> Ligne 19 Fac = 1.04 Qualité = 70.06 Contribution = 5.96**

**TI - MOULDING AND PROCESS PROPERTY RELATIONSHIPS OF A WHOLLY AROMATIC THERMOTROPIC LIQUID CRYSTAL POLYMER**

**SO - Brookfield Center, Ct., 1986, p.153-77. 012**

**Résumé Document 00000717 ==> Ligne 20 Fac = 1.03 Qualité = 71.06 Contribution = 8.50**

**TI - NOVEL INJECTION MOULDING TECHNIQUES MOVE TOWARD COMMERCIALISATION**

**SO - Plast.Technol.; 38, No.11, Oct.1992, p.17/9**

6. Les stratifiés ("laminates")

**FACTEUR 5 Coordonnees Negatives**

**Résumé Document 00000571 ==> Ligne 35 Fac = -2.65 Qualité = 286.19 Contribution = 76.99**

**TI - SELF-REINFORCED THERMOPLASTIC-LCP PREPREGS AND LAMINATES**

**SO - Polym.Composites; 15, No.4, Aug.1944, p.254-60**

**Résumé Document 00000608 ==> Ligne 41 Fac = -3.82 Qualité = 274.31 Contribution = 126.59**

**TI - SELF-REINFORCED LCP-LCP PREPREGS AND LAMINATES**

**SO - New Orleans, La., 9th-13th May 1993, Vol.II, p.1176-82. 012**

**Résumé Document 00000663 ==> Ligne 25 Fac = -0.90 Qualité = 73.51 Contribution = 4.90**

**TI - LIQUID CRYSTALLINE POLYMER/FLUOROPOLYMER BLENDS: PREPARATION AND PROPERTIES OF UNIDIRECTIONAL PREPREGS AND COMPOSITE LAMINATES**

**SO - Polym.Engng.Sci.; 33, No.13, Mid-July 1993, p.838-44**

**Résumé Document 00000769 ==> Ligne 31 Fac = -2.10 Qualité = 92.13 Contribution = 19.94**

**TI - BLENDS CONTAINING LIQUID CRYSTALLINE POLYMERS:  
PREPARATION AND PROPERTIES OF UNIDIRECTIONAL 'PREPREGS' AND  
LAMINATED COMPOSITES**

**SO - Montreal, 5th-9th May 1991, p.924-7. 012**

**Résumé Document 00000720 ==> Ligne 34 Fac = -2.58 Qualité = 75.14 Contribution = 35.09**

**TI - MESOMORPHIC AND FERROELECTRIC PROPERTIES OF  
FERROELECTRIC LIQUID CRYSTAL POLYMER/FERROELECTRIC LIQUID  
CRYSTAL BINARY MIXTURES**

**SO - Polym.Adv.Technol.; 3, No.5, Aug.1992, p.197-203**

**Résumé Document 00000984 ==> Ligne 39 Fac = -3.44 Qualité = 134.27 Contribution = 53.58**

**TI - PURITY OF LIQUID CRYSTAL POLYSILOXANES. I. ISOLATION  
PROCEDURES**

**SO - Makromol.Chem.; 188, No.11, Nov.1987, p.2759-67**

**Résumé Document 00000597 ==> Ligne 40 Fac = -3.55 Qualité = 173.41 Contribution = 61.80**

**TI - SELF-REINFORCED LIQUID CRYSTALLINE POLYMER(LCP)-LCP  
PREPREGS AND LAMINATES**

**SO - Akron, Oh., 5th-6th Oct.1993, p.339-68. 6125**

7. Les LCP utilisés comme films barrière

**FACTEUR 7 Coordonnées Positives au niveau de qualité : [ 67 , 250 ]**

**Résumé Document 00000605 ==> Ligne 4 Fac = 1.59 Qualité = 111.78 Contribution = 22.77**

**TI - PETP-LCP FILM MOVING INTO HIGH-BARRIER MARKET**

**SO - Plast.World; 52, No.5, May 1994, p.18**

**Résumé Document 00000883 ==> Ligne 5 Fac = 1.56 Qualité = 72.51 Contribution = 8.52**

**TI - BIAXIALLY STRETCHED LCP**

**SO - Plast.Ind.News(Jap.); 36, No.5, May 1990, p.68-9**

**Résumé Document 00000588 ==> Ligne 7 Fac = 1.08 Qualité = 77.61 Contribution = 9.18**

**TI - COEXTRUSION MAKES LCP COMPETITIVE IN BARRIER PACKAGING**

**SO - Plast.Technol.; 40, No.7, July 1994, p.104**

**Résumé Document 00000611 ==> Ligne 7 Fac = 1.17 Qualité = 77.38 Contribution = 11.31**

**TI - LCPS USED IN HIGH BARRIER CO-EX FILM**

**SO - Eur.Plast.News; 21, No.5, May 1994, p.18**

8. Organisation en lamelles des films extrudés.

**Résumé N0 00000587**

**TI - MORPHOLOGY OF A THERMOTROPIC RANDOM TERPOLYMER LIQUID CRYSTAL POLYMER CRYSTALLISED IN THE BULK. COMPRESSION MOULDINGS, EXTRUDATES AND FIBRES**

**AU - Kent S L; Rybnikar F; Geil P H; Carter J D ;**

**SO - Polymer; 35, No.9, 1994, p.1869-79**

**DT - Journal**

**LA - English**

**AB - The morphology of extruded pellets, compression and injection moulded articles, and melt spun fibres of a random liquid crystal terpolymer was investigated. Etching and self-decoration techniques were used and they revealed a lamellar morphology in bulk samples. The lamellae, ranging from 50 to 1800Å in thickness were oriented perpendicular to the fibre or extrusion direction in melt spun fibres and extruded pellets. Extruded pellets were analogous to a composite material with 0.2-1.8 micrometre diameter fibres of the liquid crystal being embedded in a less oriented matrix. Greater alignment of the lamellae was observed in the fibres. The thickness of the lamellae observed in melt spun fibres ranged from 500 to 800Å. Thinner lamellae (50-130Å) were observed on the surface of compression moulded specimens with thicker (200-1800Å), presumably extended chain lamellae in the interior. The thickness and degree of alignment, or order, of the lamellae was dependent on the cooling rate. Selected area electron diffraction confirmed that the chain axes were perpendicular to the lamellae. The material was a nearly equimolar random terpolymer of azelaic acid, hydroxybenzoic acid and hydroquinone. 24 . (Presented at International Polymer Physics Symposium Honouring Prof. John D Hoffman's 70th birthday, 15th-16th May 1993, Washington, DC, USA).**

**PY - 1994**

9. Etudes thermiques :

Nous notons que tous ces documents ont comme descripteur *DIFFERENTIAL THERMAL ANALYSIS* et *DSC* (DSC pour Differential Scanning Calorimetry)

**FACTEUR 10 Coordonnees Positives au niveau de qualité : [ 67 , 250 [**

**Résumé Document 00000825 ==> Ligne 9 Fac = 2.94 Qualité = 79.86 Contribution = 50.38**

**TI - ISOTHERMAL HEAT TREATMENT OF A THERMOTROPIC LCP FIBRE  
SO - J.Polym.Sci.Polym.Phys.; 29, No.4, 30th March 1991, p.395-405**

**AB - The effect of heat treatment on as-spun fibres made of a commercial liquid crystalline polymer (Vectra A900) was studied by tensile measurements and DSC. The heat treatment increased fibre strength considerably and also elongation at break but had a weaker effect on modulus. DSC showed melting behaviour to be strongly dependent on treatment temp., the melting curve displaying two endotherms whose temperature increased with treatment time.**

**DE - COPOLYESTER ; CRYSTALLISATION ; DATA; ELONGATION AT BREAK;  
ENDOTHERM ; FIBRE ; GRAPH ; HEAT TREATMENT ; HYDROXYBENZOIC  
ACID COPOLYMER ; HYDROXYNAPHTHOIC ACID COPOLYMER ;  
ISOTHERMAL ; LIQUID CRYSTAL ; MECHANICAL PROPERTIES; MELTING ;  
PLASTIC ; TECHNICAL ; TEMPERATURE DEPENDENCE ; TENSILE  
MODULUS ; TENSILE PROPERTIES ; TENSILE STRENGTH ; TEST ; TESTING ;  
THERMAL PROPERTIES; THERMOTROPIC  
PY - 1991**

**Résumé Document 00000848 ==> Ligne 11 Fac = 2.67 Qualité = 88.54 Contribution = 41.57**

**TI - MELTING BEHAVIOUR OF A THERMOTROPIC LCP (VECTRA A900) FIBRE  
SO - Polym.Comm.; 31, No.10, Oct.1990, p.372-4**

**AB - The melting behaviour of thermotropic liquid crystalline polymer as-spun fibres was investigated by DSC. The fibres displayed a double melting peak with an exotherm occurring close to the lower temperature peak. This exotherm and double melting peak were not discernible in the second heating run, indicating that they were closely connected to the thermal and stress history of the sample.**

**Résumé Document 00000608 ==> Ligne 15 Fac = 2.02 Qualité = 76.34 Contribution = 39.03**

**TI - SELF-REINFORCED LCP-LCP PREPREGS AND LAMINATES**

**SO - New Orleans, La., 9th-13th May 1993, Vol.II, p.1176-82. 012**

**AB - Self-reinforced prepregs were prepared by stretching sheets made of blends of two liquid crystal copolyesters, i.e. Vectra A950 (Hoechst Celanese) and Ultrax KR 4003 (BASF), exiting from a coathanger die attached to a static mixer/single-screw extruder set-up. Processing at this stage was carried out at temperatures at which both components were melt processable. The prepregs were laid up in multi-layers and compression moulded into quasi-isotropic or unidirectional laminates at temperatures below the melt processing temperature of one component but above the melting point of the other. The mechanical properties of the prepregs and laminates were measured. The properties of quasi-isotropic laminates were calculated based on those of unidirectional laminates and compared with measured properties. This led to the conclusion that some damage occurred during consolidation of quasi-isotropic laminates. 14 .**

**Résumé Document 00000504 ==> Ligne 17 Fac = 1.74 Qualité = 71.47 Contribution = 31.76**

**TI - PRELIMINARY INVESTIGATION OF FLASH FORMATION DURING INJECTION MOULDING OF POLYPHENYLENE SULPHIDE(PPS) AND LIQUID CRYSTALLINE POLYMER(LCP) BLENDS**

**SO - Polym.Engng.Sci.; 35, No.3, Mid-Feb.1995, p.240-6**

**AB - Flashing in pure PPS and a blend containing PPS and LCP during injection moulding was studied by DSC and SEM. The shape of the flash was observed by use of a projector. Flashing was detected in pure PPS and 90/10 PPS/LCP blend but was not found in other compositions, including pure LCP. The DSC thermograms of the flash revealed both exothermic and endothermic peaks at around 120 and 285C. The first peak, known as crystallisation temp. on heating, occurred as a result of early crystallisation of PPS. The observed double peaks indicated that the degree of crystallinity was lower in the flash than in the moulded part. The morphological studies revealed the presence of LCP fibrils in the skin region and droplets in the core region of 90/10 PPS blend. The absence of flash was attributed to the diameters of the fibrils and droplets, which were found to increase with increasing LCP component. 12 .**

**Résumé Document 00000797 ==> Ligne 17 Fac = 1.75 Qualité = 82.75 Contribution = 14.01**

**TI - CRYSTALLISATION BEHAVIOUR OF POLYPHENYLENE SULPHIDE**

**SO - Polym.Engng.Sci.; 31, No.17, Mid-Sept.1991, p.1271-8**

**AB - The crystallisation behaviour of polyphenylene sulphide (PPS) has been examined by DSC as a function of melt temperature, residence time in the melt, and the presence of a liquid crystalline polymer. Thermal history was found to play a significant part in the determination of crystalline kinetics. Short residence times at low melt temperatures and high melt temperatures alone resulted in a low value for the Avrami exponent. The former effect was ascribed to incomplete melting of the polymer while the latter effect was attributed to thermal degradation. Blending a liquid crystal polymer, Vectra A950, with PPS had only a minor effect on the crystallisation behaviour. 46 .**

## RESULTATS N° 12

**Analyse factorielle de correspondance des documents n'apparaissant pas sur les 30 premiers facteurs de deux analyses précédentes (analyse de second niveau sur les documents perdus au premier niveau).**

SUJET Polymères à cristaux liquides (LCP).

Liste des titres des documents montrant que le thème "structure du LCP" est centré sur les propriétés physiques (optiques, électriques, électroniques, ferroélectriques) obtenues à partir de l'organisation des chaînes LCP alors que les analyses précédentes portaient sur les techniques physiques permettant l'analyse de la structure du LCP

A ce thème "propriétés physiques" s'ajoute celui de la synthèse de nouveau de nouveaux LCP, qui reste toutefois très modestement représenté dans cette analyse comme dans la base entière

*FACTEUR 2 Coordonnées Positives au niveau de qualité [ 250 , 500 [*

Résumé Document 00000617 ==> Ligne 10 Fac = 2 57 Qualité = 265 00 Contribution = 63 88

**TI - NOVEL OPTICAL METHOD FOR MONITORING HOMEOTROPIC ALIGNMENT, APPLIED TO A SIDE CHAIN LIQUID CRYSTAL POLYMER IN AC AND DC ELECTRIC FIELDS**

SO - Molecular Crystals è Liquid Crystals; Vol.241, 1994, p.255-73

*FACTEUR 2 Coordonnées Positives au niveau de qualité . [ 67 , 250 [*

Résumé Document 00000630 ==> Ligne 4 Fac = 3 13 Qualité = 243 54 Contribution = 39 79

**TI - DESIGN AND SYNTHESIS OF A NEW SIDE-CHAIN LIQUID CRYSTALLINE POLYMER BY METATHESIS POLYMERISATION**

SO - Macromolecules; 27, No.1, 3rd Jan.1994, p.309-11

Résumé Document 00000818 ==> Ligne 7 Fac = 2.88 Qualité = 83.71 Contribution = 33.60

**TI - POLYMER SYNTHESIS PRODUCES LIQUID CRYSTAL POLYMER**

SO - New Mat./Jap.; Aug.1991, p.12-3

Résumé Document 00000679 ==> Ligne 8 Fac = 2.74 Qualité = 101.80 Contribution = 34.33

**TI - STUDY OF A CHOLESTERIC LIQUID CRYSTAL POLYMER IN SOLUTION IN A LOW MOLAR WEIGHT LIQUID CRYSTAL: PHASE DIAGRAM AND ELECTRIC FIELD EFFECT**

SO - Molecular Crystals è Liquid Crystals; Vol.231, 1993, p.11-28

Résumé Document 00000635 ==> Ligne 9 Fac = 2.59 Qualité = 220.08 Contribution = 67.90

**TI - DIELECTRIC RELAXATIONS OF A SMECTIC SIDE-CHAIN LIQUID-CRYSTALLINE POLYMER IN DIFFERENT ALIGNMENT STATES**

SO - Molecular Crystals è Liquid Crystals; Vol.238, 1994, p.129-45

**Résumé Document 00001036 ==> Ligne 9 Fac = 2.61 Qualité = 154.41 Contribution = 27.68**

**TI - ELECTRO-OPTIC EFFECTS IN A SMECTOGENIC POLYSILOXANE SIDE-CHAIN LIQUID CRYSTAL POLYMER**

**SO - Lyngby, Denmark, Sept.12-14, 1983, p.323-34. 9112**

**Résumé Document 00000660 ==> Ligne 11 Fac = 2.47 Qualité = 169.26 Contribution = 52.44**

**TI - MAGNETIC ORIENTATION AND MICROSTRUCTURE OF MAIN-CHAIN THERMOTROPIC COPOLYESTERS**

**SO - Polymer; 34, No.16, 1993, p.3347-57**

**Résumé Document 00000787 ==> Ligne 11 Fac = 2.42 Qualité = 83.00 Contribution = 26.81**

**TI - CONVENIENT SYNTHESIS OF A LIQUID CRYSTALLINE POLYACRYLATE WITH A TETRAMETHYLENE SPACER**

**SO - Polym.Comm.; 32, No.12, 1991, p.380-1**

**Résumé Document 00000682 ==> Ligne 14 Fac = 2.16 Qualité = 135.36 Contribution = 21.25**

**TI - THERMO-OPTIC OBSERVATIONS ON A MAIN-CHAIN THERMOTROPIC LIQUID CRYSTAL POLYMER**

**SO - Polym.Bull.; 30, No.5, May 1993, p.595-601**

**Résumé Document 00001029 ==> Ligne 15 Fac = 2.07 Qualité = 148.54 Contribution = 30.34**

**TI - SURFACE-ENHANCED RATE OF MOLECULAR ALIGNMENT IN A LIQUID CRYSTAL POLYMER**

**SO - Macromolecules; 19, No.9, Sept.1986, p.2459-61**

**Résumé Document 00001064 ==> Ligne 15 Fac = 2.07 Qualité = 187.85 Contribution = 34.91**

**TI - SOME ELECTRO-OPTICAL PHENOMENA IN COMB-LIKE LIQUID CRYSTALLINE POLYMERIC AZOMETHYNES**

**SO - Polym.Preprints; 24, No.2, Aug.1983, p.309-10**

**Résumé Document 00000544 ==> Ligne 16 Fac = 1.98 Qualité = 130.21 Contribution = 15.86**

**TI - FIBRES FROM POLYPROPYLENE AND LIQUID CRYSTAL POLYMER BLENDS USING COMPATIBILISING AGENTS. II. SYNTHESIS OF GRAFT SIDE CHAIN LIQUID CRYSTALLINE POLYMERS BASED UPON ACRYLIC ACID-FUNCTIONALISED POLYPROPYLENE**

**• SO - Macromolecular Rapid Communications; 15, No.11, Nov.1994, p.857-65**

**Résumé Document 00000697 ==> Ligne 16 Fac = 1.97 Qualité = 127.96 Contribution = 19.65**

**TI - INVESTIGATION OF MELT DIFFUSION IN LIQUID CRYSTALLINE POLYMERS BY FORWARD RECOIL SPECTROMETRY (FRES)**

**SO - Chimica e Ind.; 74, Nos.8/9, Aug./Sept.1992, p.579-84**

**Résumé Document 00000795 ==> Ligne 17 Fac = 1.86 Qualité = 89.02 Contribution = 12.27**

**TI - FOLDED CHAIN SINGLE CRYSTALS OF A LIQUID CRYSTAL POLYMER  
SO - J.Mat.Sci.Lett.; 10, No.24, 15th Dec.1991, p.1456-60**

**Résumé Document 00000945 ==> Ligne 18 Fac = 1.80 Qualité = 117.42 Contribution = 16.36**

**TI - LASER-INDUCED PHOTO-OPTICAL RECORDING ON FREE-STANDING FILMS OF A MAIN-CHAIN NEMATIC POLYESTER  
SO - Makromol.Chem.Rapid Commun.; 9, No.7, July 1988, p.463-7**

**Résumé Document 00000918 ==> Ligne 19 Fac = 1.62 Qualité = 154.56 Contribution = 20.00**

**TI - SELF-REINFORCING POLYMERS  
SO - Hoechst High Chem Mag.; No.6, 1989, p.36-41**

**Résumé Document 00000712 ==> Ligne 21 Fac = 1.45 Qualité = 86.29 Contribution = 15.02**

**TI - COMMUNICATIONS. FAST ELECTROCLINIC SWITCHING IN A FERROELECTRIC LC-POLYSILOXANE  
SO - Adv.Mat.; 4, No.12, 1992, p.792-4**

## RESULTATS N°13

### Analyse factorielle de correspondance des documents n'apparaissant pas sur les 30 premiers facteurs de deux analyses précédentes

Mise en évidence de l'orientation "composite" du thème "mélange" ("blends") dans cette analyse

*FACTEUR 1 Coordonnées Positives au niveau de qualité · [ 67 , 250 [*

**Résumé** Document 00000533 ==> Ligne 4 Fac = 1 36 Qualité = 84.94 Contribution = 23.58

**TI - CRITICAL BEHAVIOUR OF SHEAR-INDUCED TRANSIENT PERIODIC STRUCTURES IN A LYOTROPIC LIQUID CRYSTALLINE POLYMER AS A FUNCTION OF MOLECULAR WEIGHT**

SO - *Macromolecules*; 27, No.26, 19th Dec.1994, p.7843-5

**Résumé** Document 00000855 ==> Ligne 6 Fac = 1 22 Qualité = 130 18 Contribution = 6 15

**TI - NOVEL REINFORCED POLYMERS BASED ON BLENDS OF POLYSTYRENE AND A THERMOTROPIC LIQUID CRYSTALLINE POLYMER**

SO - Rosemont, Il., 23rd-24th Sept.1987, p.267-76. 63E

**Résumé** Document 00000918 ==> Ligne 7 Fac = 1 13 Qualité = 75 49 Contribution = 7 92

**TI - SELF-REINFORCING POLYMERS**

SO - *Hoechst High Chem Mag.*; No.6, 1989, p.36-41

**Résumé** Document 00001035 ==> Ligne 7 Fac = 1 16 Qualité = 110 32 Contribution = 5 52

**TI - SELF REINFORCING POLYMER BLENDS OF THERMOTROPIC LIQUID CRYSTAL POLYMER AND POLYSTYRENE**

SO - Boston, April 28-May 1, 1986, p.306-10. 012

**Résumé** Document 00000754 ==> Ligne 9 Fac = 0 99 Qualité = 71 35 Contribution = 6.81

**TI - SHEAR FLOW CHARACTERISATION OF BLENDS CONTAINING LIQUID CRYSTAL POLYMERS**

SO - *Makromol.Chem.,Macromol.Symp.*; 56, April 1992, p.151-9

**Résumé** Document 00000640 ==> Ligne 10 Fac = 0 96 Qualité = 72 38 Contribution = 2 28

**TI - WHOLLY AROMATIC THERMOTROPIC LIQUID CRYSTALLINE COPOLYESTERS AND THEIR IN-SITU COMPOSITES**

SO - Madrid, 12th-16th July 1993, p.679-83. 627

**Résumé** Document 00000715 ==> Ligne 10 Fac = 0 91 Qualité = 126.83 Contribution = 5 44

**TI - RHEOLOGY OF RIGID-FLEXIBLE LIQUID CRYSTALLINE POLYMER (LCP)/PMMA BLENDS**

SO - *J.Appl.Polym.Sci.*; 46, No.6, 25th Oct.1992, p.959-63

**Résumé** Document 00000698 ==> Ligne 14 Fac = 0 66 Qualité = 86.30 Contribution = 3 38

**TI - STUDIES ON BLENDS OF A THERMOTROPIC LIQUID CRYSTALLINE POLYMER AND POLYBUTYLENE TEREPHTHALATE**

SO - *Polym.Bull.*; 30, No.3, March 1993, p.353-60

**RESULTATS N°14**  
**EVOLUTION CHRONOLOGIQUE DES THEMES D'UNE BASE DE DONNEES**  
**Base RAPRA, sujet CAPTEURS**  
**Contenu des classes de mots**

**classe 1**

NUM	REFE	1	2	3	4	20	27	28	29	30
		8084	8588	8994	9597	INER	8084	8588	8994	9597
13REFE	GLUCOSE	2 0	0 0	3 0	38 0	3 1	1 0	0 0	1 0	1 0
14REFE	CHEMICAL	0 0	2 0	15 0	63 0	3 0	0 0	1 0	1 0	1 0
15REFE	POLYPYRROLE	0 0	0 0	6 0	40 0	3 0	0 0	0 0	1 0	1 0
16REFE	WATER	1 0	0 0	5 0	39 0	2 8	1 0	0 0	1 0	1 0
18REFE	EMBEDDED	0 0	0 0	5 0	36 0	2 8	0 0	0 0	1 0	1 0
19REFE	FOAM	4 0	0 0	1 0	26 0	2 6	1 0	0 0	1 0	1 0
22REFE	ENZYME	0 0	0 0	2 0	27 0	2 6	0 0	0 0	1 0	1 0
23REFE	ION	0 0	0 0	7 0	38 0	2 5	0 0	0 0	1 0	1 0
24REFE	ELECTRODE	0 0	1 0	3 0	32 0	2 5	0 0	1 0	1 0	1 0
25REFE	SUBSTRATE	3 0	0 0	1 0	27 0	2 5	1 0	0 0	1 0	1 0
26REFE	CONDUCTING	0 0	0 0	9 0	41 0	2 5	0 0	0 0	1 0	1 0
31REFE	MATRIX	0 0	0 0	13 0	46 0	2 3	0 0	0 0	1 0	1 0
32REFE	SENSITIVITY	0 0	0 0	4 0	30 0	2 3	0 0	0 0	1 0	1 0
33REFE	INC	0 0	0 0	6 0	34 0	2 3	0 0	0 0	1 0	1 0
34REFE	LIQUID	0 0	2 0	3 0	32 0	2 3	0 0	1 0	1 0	1 0
37REFE	STRAIN	0 0	6 0	7 0	44 0	2 3	0 0	1 0	1 0	1 0
39REFE	SILICONE	1 0	0 0	2 0	25 0	2 0	1 0	0 0	1 0	1 0
41REFE	ELEMENT	0 0	1 0	2 0	25 0	2 0	0 0	1 0	1 0	1 0
45REFE	POTENTIAL	1 0	0 0	6 0	33 0	2 0	1 0	0 0	1 0	1 0
46REFE	VALUE	2 0	0 0	5 0	32 0	1 9	1 0	0 0	1 0	1 0
47REFE	RAM	0 0	3 0	1 0	23 0	1 9	0 0	1 0	1 0	1 0
49REFE	POLYMERIC	0 0	1 0	6 0	32 0	1 8	0 0	1 0	1 0	1 0
50REFE	SOLUTION	0 0	1 0	5 0	30 0	1 8	0 0	1 0	1 0	1 0
51REFE	VAPOUR	0 0	1 0	5 0	30 0	1 8	0 0	1 0	1 0	1 0
52REFE	ORGANIC	0 0	2 0	6 0	34 0	1 8	0 0	1 0	1 0	1 0
53REFE	EXAMINED	0 0	2 0	16 0	50 0	1 8	0 0	1 0	1 0	1 0
56REFE	WAVE	0 0	1 0	12 0	41 0	1 8	0 0	1 0	1 0	1 0
57REFE	ACID	0 0	0 0	2 0	20 0	1 7	0 0	0 0	1 0	1 0
58REFE	FIXED	0 0	0 0	0 0	14 0	1 7	0 0	0 0	0 0	1 0
61REFE	INDICATOR	0 0	0 0	2 0	19 0	1 6	0 0	0 0	1 0	1 0
62REFE	OXIDASE	0 0	0 0	2 0	19 0	1 6	0 0	0 0	1 0	1 0
63REFE	HAVING	0 0	0 0	3 0	21 0	1 6	0 0	0 0	1 0	1 0
66REFE	SO	1 0	0 0	7 0	31 0	1 6	1 0	0 0	1 0	1 0
67REFE	UNIFORM	0 0	0 0	0 0	13 0	1 6	0 0	0 0	0 0	1 0
68REFE	PRODUCES	0 0	0 0	0 0	13 0	1 6	0 0	0 0	0 0	1 0
69REFE	MEMBRANES	0 0	0 0	9 0	31 0	1 6	0 0	0 0	1 0	1 0
70REFE	HYDRAULIC	1 0	0 0	5 0	27 0	1 5	1 0	0 0	1 0	1 0
72REFE	THIN	0 0	2 0	13 0	42 0	1 5	0 0	1 0	1 0	1 0
73REFE	HEATER	1 0	0 0	2 0	20 0	1 5	1 0	0 0	1 0	1 0
75REFE	ACTUATOR	0 0	0 0	3 0	20 0	1 5	0 0	0 0	1 0	1 0
76REFE	ELECTROCHEMICAL	0 0	0 0	4 0	22 0	1 5	0 0	0 0	1 0	1 0
78REFE	POLY	0 0	2 0	3 0	24 0	1 5	0 0	1 0	1 0	1 0
80REFE	FLUORESCENT	0 0	0 0	1 0	15 0	1 5	0 0	0 0	1 0	1 0
82REFE	SOLUTIONS	0 0	0 0	2 0	17 0	1 4	0 0	0 0	1 0	1 0
83REFE	METAL	1 0	0 0	9 0	32 0	1 4	1 0	0 0	1 0	1 0
84REFE	SAMPLE	0 0	0 0	4 0	21 0	1 4	0 0	0 0	1 0	1 0
85REFE	SHOWED	0 0	0 0	6 0	24 0	1 3	0 0	0 0	1 0	1 0
86REFE	PHASE	0 0	1 0	3 0	21 0	1 3	0 0	1 0	1 0	1 0
89REFE	CHANGE	0 0	0 0	7 0	25 0	1 3	0 0	0 0	1 0	1 0
90REFE	DIOXIDE	0 0	0 0	4 0	20 0	1 3	0 0	0 0	1 0	1 0

## RESULTATS N°14 (suite)

### classe 2

NUM	REFE	1	2	3	4	20	27	28	29	30
		8084	8588	8994	9597	INER	8084	8588	8994	9597
17	REFE COMPOSITE	0 0	1 0	49 0	79 0	2 8	0 0	1 0	1 0	1 0
87	REFE FLUORESCENCE	0 0	0 0	21 0	27 0	1 3	0 0	0 0	1 0	1 0
94	REFE RHEOMETER	0 0	0 0	14 0	12 0	1 1	0 0	0 0	1 0	1 0
95	REFE REPORTED	0 0	7 0	36 0	47 0	1 0	0 0	1 0	1 0	1 0
97	REFE SHOWN	0 0	2 0	21 0	22 0	1 0	0 0	1 0	1 0	1 0
100	REFE GRAPHITE	0 0	0 0	12 0	11 0	0 9	0 0	0 0	1 0	1 0
101	REFE SHEAR	0 0	0 0	12 0	11 0	0 9	0 0	0 0	1 0	1 0
106	REFE PROFILES	0 0	1 0	15 0	16 0	0 8	0 0	1 0	1 0	1 0

### classe 3

NUM	REFE	1	2	3	4	20	27	28	29	30
		8084	8588	8994	9597	INER	8084	8588	8994	9597
4	REFE SITU	0 0	0 0	52 0	20 0	8 0	0 0	0 0	1 0	1 0
5	REFE CURE	0 0	13 0	99 0	65 0	7 5	0 0	1 0	1 0	1 0
7	REFE VISCOSITY	0 0	14 0	43 0	15 0	6 5	0 0	1 0	1 0	1 0
10	REFE DEPENDENT	0 0	1 0	23 0	4 0	4 7	0 0	1 0	1 0	1 0
11	REFE ELECTROMAGNETIC	2 0	0 0	22 0	3 0	4 6	1 0	0 0	1 0	1 0
12	REFE EMPHASIS	1 0	5 0	7 0	0 0	3 1	1 0	1 0	1 0	0 0
20	REFE TG	0 0	5 0	9 0	2 0	2 6	0 0	1 0	1 0	1 0
21	REFE HDPE	0 0	2 0	16 0	4 0	2 6	0 0	1 0	1 0	1 0
27	REFE REACTION	0 0	6 0	29 0	17 0	2 4	0 0	1 0	1 0	1 0
29	REFE COMPOSITES	0 0	5 0	46 0	42 0	2 4	0 0	1 0	1 0	1 0
35	REFE LASER	1 0	0 0	24 0	15 0	2 3	1 0	0 0	1 0	1 0
36	REFE CONTAINERS	0 0	0 0	13 0	4 0	2 3	0 0	0 0	1 0	1 0
38	REFE SPC	0 0	3 0	11 0	2 0	2 1	0 0	1 0	1 0	1 0
42	REFE FABRICATION	1 0	0 0	16 0	7 0	2 0	1 0	0 0	1 0	1 0
43	REFE CURING	0 0	4 0	32 0	25 0	2 0	0 0	1 0	1 0	1 0
44	REFE ASSURANCE	1 0	3 0	7 0	0 0	2 0	1 0	1 0	1 0	0 0
48	REFE AUTOMATED	0 0	2 0	18 0	9 0	1 8	0 0	1 0	1 0	1 0
54	REFE PROCESSOR	1 0	0 0	14 0	6 0	1 8	1 0	0 0	1 0	1 0
55	REFE MOVING	0 0	0 0	12 0	5 0	1 8	0 0	0 0	1 0	1 0
64	REFE RTM	0 0	2 0	14 0	6 0	1 6	0 0	1 0	1 0	1 0
74	REFE SEALING	0 0	1 0	10 0	3 0	1 5	0 0	1 0	1 0	1 0
77	REFE DYNISCO	0 0	1 0	12 0	5 0	1 5	0 0	1 0	1 0	1 0
79	REFE ADVANCED	0 0	3 0	19 0	12 0	1 5	0 0	1 0	1 0	1 0
81	REFE NOZZLE	0 0	2 0	19 0	13 0	1 4	0 0	1 0	1 0	1 0
92	REFE AUTOCLAVE	0 0	1 0	9 0	3 0	1 2	0 0	1 0	1 0	1 0
96	REFE THERMOCOUPLE	0 0	2 0	11 0	6 0	1 0	0 0	1 0	1 0	1 0
98	REFE RHEOLOGICAL	0 0	0 0	10 0	7 0	1 0	0 0	0 0	1 0	1 0
99	REFE REVIEWED	0 0	3 0	18 0	15 0	1 0	0 0	1 0	1 0	1 0
102	REFE CORRELATED	0 0	0 0	9 0	6 0	0 9	0 0	0 0	1 0	1 0
103	REFE FLOWING	0 0	0 0	8 0	5 0	0 9	0 0	0 0	1 0	1 0
104	REFE RECYCLING	0 0	0 0	8 0	5 0	0 9	0 0	0 0	1 0	1 0
105	REFE ENGINEERING	0 0	1 0	13 0	11 0	0 8	0 0	1 0	1 0	1 0

### classe 5 et 4

NUM	REFE	1	2	3	4	20	27	28	29	30
		8084	8588	8994	9597	INER	8084	8588	8994	9597
1	REFE VISION	0 0	10 0	1 0	2 0	10 6	0 0	1 0	1 0	1 0
2	REFE PRODUCTIVITY	9 0	3 0	0 0	7 0	9 2	1 0	1 0	0 0	1 0
3	REFE SCANNER	7 0	1 0	4 0	0 0	9 1	1 0	1 0	1 0	0 0
6	REFE ANGLE	7 0	0 0	0 0	9 0	6 6	1 0	0 0	0 0	1 0
8	REFE PATTERN	6 0	0 0	1 0	8 0	4 9	1 0	0 0	1 0	1 0
9	REFE INFRA	5 0	0 0	4 0	2 0	4 8	1 0	0 0	1 0	1 0
28	REFE PROTECTION	1 0	5 0	0 0	7 0	2 4	1 0	1 0	0 0	1 0
30	REFE BETA	4 0	0 0	1 0	9 0	2 4	1 0	0 0	1 0	1 0
40	REFE INCORPORATED	4 0	0 0	2 0	9 0	2 0	1 0	0 0	1 0	1 0
59	REFE PRE	3 0	1 0	0 0	8 0	1 7	1 0	1 0	0 0	1 0
60	REFE BUILT	3 0	0 0	2 0	5 0	1 6	1 0	0 0	1 0	1 0
65	REFE CONSISTS	3 0	2 0	0 0	9 0	1 6	1 0	1 0	0 0	1 0
71	REFE ALLOW	3 0	0 0	1 0	10 0	1 5	1 0	0 0	1 0	1 0
88	REFE CONSISTING	3 0	0 0	2 0	8 0	1 3	1 0	0 0	1 0	1 0
91	REFE CHEMICALS	2 0	0 0	0 0	8 0	1 2	1 0	0 0	0 0	1 0
93	REFE ORIENTATION	3 0	0 0	2 0	10 0	1 2	1 0	0 0	1 0	1 0

## RESULTATS N° 15

Désambiguïsation sans dictionnaire

Cas de l'analyse CAPTEURS: comment le thème des propriétés des polymères pour être capteurs s'isole de celui des capteurs utilisés dans les procédés de transformation des polymères

THEME : CAPTEURS ET POLYMERE

OPPOSITION ENTRE LES MATERIAUX POLYMERES ENTRANT DANS LA COMPOSITION DE CAPTEURS ET LES CAPTEURS UTILISES DANS LA TRANSFORMATION DES POLYMERES POUR CONTROLER LE PROCESS

### CAPTEURS (ANALYSE 3)

7 mots et 18 documents, associés sur ce facteur 4 positif, se réfèrent aux matériaux polymères entrant dans la composition de capteurs

IL Y A 7 MOTS EXPLICATIFS :

POUR LE FACTEUR 4 LISTE DES MOTS A COORDONNEE POSITIVE :

LIGNE	FACTEUR	QUALITE	CONTRI		
1	3.61	319 78	204 85	86	POLYMERS
22	1.19	98 78	41 09	4	APPLICATIONS
22	1.23	64 63	38 14	77	OPTICAL
22	1.26	66 04	26 98	127	TEMPERATURES
24	1.03	89 17	36 81	85	POLYMER
25	0.92	104.12	44 54	67	MELT
29	0.49	58 00	26 22	126	TEMPERATURE

QUALITE = 250' ( 1 ) doc

QUALITE = 67' ( 17 ) doc

Titre des documents :

FACTEUR 4 Coordonnees Positives au niveau de qualité [ 67 , 250 [

Résumé Document 00000287 ==> Ligne 1 Fac = 3.56 Qualité = 211.88 Contribution = 36.92

@TI DESIGN AND FABRICATION OF PERM SELECTIVE MEMBRANES FOR ISFET SENSORS

@SO Canterbury, 11th-12th Sept.1990, p.P7/1-P7/5. 98

Résumé Document 00000280 ==> Ligne 9 Fac = 2.65 Qualité = 200.01 Contribution = 32.73

@TI ELECTRICALLY CONDUCTING POLYMERS

@SO Mat.Des.; 11, No.3, June 1990, p.142-52

**Résumé Document 00000132 ==> Ligne 13 Fac = 2.27 Qualité = 190.20 Contribution = 17.93**

**@TI CONDUCTIVE POLYMERS: EVALUATION OF INDUSTRIAL APPLICATIONS**

**@SO Synthetic Metals; 57, No.1, 12th April 1993, p.3623-31**

**Résumé Document 00000258 ==> Ligne 16 Fac = 1.95 Qualité = 205.67 Contribution = 15.43**

**@TI FUNCTIONALISATION OF CONDUCTING POLYMER MEMBRANES**

**@SO Chimica e Ind.; 73, No.12, Dec.1991, p.960-70**

**Résumé Document 00000061 ==> Ligne 17 Fac = 1.79 Qualité = 200.45 Contribution = 31.62**

**@TI ACS POLYMERIC MATERIALS SCIENCE & ENGINEERING SPRING MEETING 1994. VOLUME 70. Proceedings of a conference held San Diego, Ca., 13th-18th March 1994**

**@SO Washington, D.C., 1993, pp.xxxviii, 520. 11ins. 28/3/94. NALOAN.**

**Résumé Document 00000415 ==> Ligne 18 Fac = 1.74 Qualité = 179.06 Contribution = 44.05**

**@TI OPTICAL POLYMERS STUDY FINDS NEW APPLICATIONS**

**@SO Plast.Rubb.Wkly.; No.1147, 19th July 1986, p.6**

**Résumé Document 00000149 ==> Ligne 21 Fac = 1.35 Qualité = 131.34 Contribution = 12.72**

**@TI PORE-SIZE-SPECIFIC MODIFICATION OF POROUS MATERIALS**

**@SO Adv.Mat.; 6, No.3, March 1994, p.242-4**

**Résumé Document 00000201 ==> Ligne 21 Fac = 1.33 Qualité = 125.07 Contribution = 12.30**

**@TI PROSPECTS FOR ELECTRONIC PROCESSES IN ORGANIC MATERIALS**

**@SO Molecular Crystals & Liquid Crystals; Vol.228/229/230, 1993, p.1-22**

**Résumé Document 00000436 ==> Ligne 22 Fac = 1.22 Qualité = 70.71 Contribution = 7.82**

**@TI MAKING THE MOST OF MELT TEMPERATURE MEASUREMENT**

**@SO Plast.Engng.; 41, No.7, July 1985, p.39-43**

**Résumé Document 00000010 ==> Ligne 23 Fac = 1.08 Qualité = 74.09 Contribution = 13.63**

**@TI MEASUREMENT OF INDIVIDUAL COEX LAYER THICKNESSES**

**@SO Brit.Plast.Rubb.; June 1995, p.19-20**

**Résumé Document 00000117 ==> Ligne 24 Fac = 1.04 Qualité = 101.96 Contribution = 14.35**

**@TI SIGNAL INTERPRETATION OF POLYMER MELT TEMPERATURES USING AN IR PYROMETER VIEWING A DEFINED THERMAL ENVIRONMENT**

**@SO New Orleans, La., 9th-13th May 1993, Vol.II, p.2218-22. 012**

**Résumé Document 00000304 ==> Ligne 24 Fac = 1.05 Qualité = 113.81 Contribution = 8.31**

**@TI INFRARED THERMOMETER**

**@SO Rubb.World; 203, No.5, Feb.1991, p.60**

**Résumé Document 00000347 ==> Ligne 25 Fac = 0.95 Qualité = 107.07 Contribution = 8.91**

**@TI DYNISCO MELT PRESSURE INSTRUMENTATION FOR THE PLASTICS EXTRUSION PROCESS**

**@SO Alton, 1989, pp.29. 11ins. 5/7/89. 28(21)-82**

**Résumé Document 00000054 ==> Ligne 26 Fac = 0.84 Qualité = 74.32 Contribution = 6.12**

**@TI TEMPERATURE PROFILES AND SHEAR HEATING EFFECTS IN POLYMER PROCESSING: TAPERED DIE**

**@SO San Francisco, Ca., 1st-5th May 1994, Vol.I, p.666-70. 012**

**Résumé Document 00000297 ==> Ligne 26 Fac = 0.76 Qualité = 77.63 Contribution = 6.68**

**@TI RHEOLOGICAL MEASUREMENTS ON THE INJECTION MOULDING MACHINE - ASPECTS OF THE DEVELOPMENT OF A PROCESS MONITORING SYSTEM**

**@SO Kunststoffe German Plastics; 81, No.3, March 1991, p.24-6**

**Résumé Document 00000453 ==> Ligne 26 Fac = 0.83 Qualité = 67.10 Contribution = 7.64**

**@TI ASSESSMENT OF PROCESSING MORPHOLOGY AND ON-LINE MEASUREMENT OF SPHERULITE SIZE FOR POLYPROPYLENE FILM**

**@SO Plast.Rubb.Process.Appln.; 4, No.3, 1984, p.247-55**

**Résumé Document 00000118 ==> Ligne 27 Fac = 0.68 Qualité = 72.97 Contribution = 4.01**

**@TI PRESSURE EFFECTS ON TEMPERATURE MEASUREMENT OF POLYMER MELT**

**@SO Polym.Bull.; 33, No.1, June 1994, p.127-32**

## RESULTATS N° 15 (suite)

### CAPTEURS MISE A JOUR (analyse 2-1)

*Dans cette version n°1 de l'analyse de la base capteurs mise à jour on retrouve isolée sur une "métaclé" (partie positive ou négative d'un facteur) l'ensemble des documents et des mots décrivant les matériaux polymères utilisés comme capteurs par opposition aux capteurs utilisés dans la transformation des polymères*

**IL Y A 80 MOTS EXPLICATIFS :**

**POUR LE FACTEUR 1 LISTE DES MOTS A COORDONNEE POSITIVE**

**SEUILS POUR LA CONTRIBUTION 2\*MOY = 2 08 3\*MOY = 3 12**

LIGNE	FACTEUR	QUALITE	CONTRI		
1	3 38	101 76	15 71	595	OXIDASE
3	3 02	83 79	10 16	28	AMPEROMETRIC
3	3 01	105 83	25 56	364	GLUCOSE
5	2 73	86 70	14 19	282	ENZYME
6	2 57	105 83	15 50	264	ELECTRODE
6	2 63	82 84	9 97	399	IMMOBILISATION
7	2 51	38 21	6 58	72	BIOSENSORS
7	2 51	25 51	4 95	646	POLYTHIOPHENE
8	2.20	53 91	4 75	400	IMMOBILISED
8	2 22	39 31	4 86	519	METHACRYLATE
8	2.22	79.90	14 78	645	POLYPYRROLE
8	2 32	46.50	6 68	729	REDOX
9	2 19	93 80	8 16	263	ELECTROCHEMICAL
9	2 10	100 48	14 72	516	MEMBRANE
9	2 04	34 01	3 82	520	METHYL
9	2 09	24 34	5 45	625	PLASMA
9	2 10	32 58	3 17	643	POLYMERISED
10	1 93	50 24	3 40	40	AQUEOUS
10	1 96	75 38	8.04	265	ELECTRODES
10	1 91	14 09	2 63	287	ETHYLENE
10	2 01	78 67	10 61	517	MEMBRANES
10	1 95	17 46	3 00	647	POROUS
10	1 90	60 37	4 23	663	PREPARED
11	1 79	31 97	4 63	7	ACID
11	1 74	22 13	2 98	244	DOPED
11	1 76	40.49	6 27	616	PH
11	1 79	33 83	6 10	635	POLY
11	1 79	40.82	3 37	935	VAPOURS
12	1 71	26 27	3 63	356	GEL
12	1 70	57 01	8 48	445	ION
12	1 66	12 26	2 88	447	IONS
12	1 64	38 70	7 92	596	OXYGEN
12	1.64	18 79	2 28	707	QUARTZ
12	1 65	33.51	3 37	753	RESPONSES
12	1 69	66 75	7 84	785	SELECTIVE
12	1 69	23.68	2 60	786	SELECTIVITY
12	1 69	17.05	2 23	831	SORPTION
13	1 42	44.38	4 07	153	CONTAINING

13	1.47	59.31	5.95	588	ORGANIC
13	1.54	17.17	2.96	614	PERMEABILITY
13	1.47	59.65	5.08	828	SOLUTION
13	1.51	11.30	2.08	861	STYRENE
14	1.32	47.55	4.53	109	COATED
14	1.37	38.32	2.57	136	CONCENTRATION
14	1.38	42.32	6.20	141	CONDUCTING
14	1.34	19.76	2.36	169	COPOLYMERS
14	1.31	20.91	2.69	236	DIOXIDE
14	1.29	81.39	10.24	322	FILMS
14	1.31	21.50	2.13	371	GROUPS
14	1.29	71.82	6.46	501	MATRIX
14	1.31	20.85	2.45	533	MODIFIED
14	1.37	54.26	4.20	790	SENSITIVITY
14	1.35	34.15	4.26	934	VAPOUR
15	1.10	30.45	3.87	8	ACOUSTIC
15	1.21	36.15	4.71	642	POLYMERISATION
15	1.25	20.84	2.94	814	SILICONE
16	1.06	56.37	5.91	100	CHEMICAL
16	0.94	24.84	2.45	111	COATINGS
16	1.08	24.39	3.66	331	FLUORESCENCE
16	1.09	16.79	2.97	392	HUMIDITY
16	0.99	48.68	7.30	644	POLYMERS
16	0.97	31.46	2.48	651	POTENTIAL
16	1.01	61.63	4.00	789	SENSITIVE
16	0.98	28.30	2.62	841	SPECTROSCOPY
16	1.06	30.81	2.62	856	STRUCTURE
16	1.05	40.63	3.92	953	WAVE
17	0.87	27.30	3.00	88	CARBON
17	0.79	42.84	4.04	283	EPOXY
17	0.88	42.22	2.58	444	INVESTIGATED
17	0.88	30.78	3.14	581	OPTIC
17	0.82	104.07	12.26	640	POLYMER
17	0.88	48.05	2.82	858	STUDIED
17	0.81	28.69	2.51	900	THIN
18	0.67	32.06	3.80	127	COMPOSITE
18	0.72	64.64	7.65	315	FIBRE
18	0.71	35.29	2.83	752	RESPONSE
19	0.59	47.42	4.09	62	BASED
19	0.59	24.99	3.58	582	OPTICAL
19	0.50	32.64	2.33	694	PROPERTIES
20	0.44	14.96	2.26	183	CURE

**QUALITE = 67(9) doc**

Analyse 2-1 : Titre des documents associé au :

**FACTEUR 1 Coordonnées Positives au niveau de qualité : [ 67 , 250 ]**

**Résumé Document 00001196 ==> Ligne 6 Fac = 2.56 Qualité = 91.14 Contribution = 15.05**

**@TI INTERFERING FACTOR IN THE GLUCOSE OXIDASE SENSING SYSTEM WITH POLYPYRROLE/GLUCOSE OXIDASE MEMBRANE**

**@SO Synthetic Metals; 71, Nos.1-3, 1st April 1995, p.2245-6**

**Résumé Document 00001022 ==> Ligne 7 Fac = 2.43 Qualité = 85.86 Contribution = 15.80**

**@TI AMPEROMETRIC BIOSENSORS FOR DETECTION OF CHOLINE AND LACTATE BY CHOLINE OXIDASE AND LACTATE OXIDASE IMMOBILISATION IN MICROPOROUS MEMBRANES**

**@SO Espoo, 1996, p.39. 30cms. 10/2/97**

**Résumé Document 00001197 ==> Ligne 7 Fac = 2.47 Qualité = 105.59 Contribution = 22.43**

**@TI ELECTROCHEMICAL FABRICATION OF AMPEROMETRIC GLUCOSE ENZYME ELECTRODE BY IMMOBILISING GLUCOSE OXIDASE IN ELECTROPOLYMERISED POLY(3,3'-DIAMINOBENZIDINE)(I) FILM ON PALLADINISED GLASSY CARBON ELECTRODE**

**@SO Analyst; 121, No.7, July 1996, p.971-5**

**Résumé Document 00001502 ==> Ligne 8 Fac = 2.24 Qualité = 122.01 Contribution = 18.66**

**@TI GLUCOSE-SENSITIVE POLYPYRROLE/POLYSTYRENESULPHONATE(PSS) FILMS CONTAINING CO-IMMOBILISED GLUCOSE OXIDASE(GOx) AND (FERROCENYLMETHYL)TRIMETHYLAMMONIUM BROMIDE(FcBr)**

**@SO J.Macromol.Sci.A; 32, No.2, 1995, p.349-59**

**Résumé Document 00001209 ==> Ligne 9 Fac = 2.16 Qualité = 79.14 Contribution = 6.74**

**@TI ELECTROCHEMICAL SYNTHESIS AND SENSOR APPLICATION OF POLYDIAMINOBENZENE**

**@SO Synthetic Metals; 79, No.1, 30th April 1996, p.57-61**

**Résumé Document 00001058 ==> Ligne 11 Fac = 1.73 Qualité = 73.31 Contribution = 9.24**

**@TI MODIFICATION OF FERROCENE-CONTAINING REDOX GEL SENSOR PERFORMANCE BY COPOLYMERISATION OF CHARGED MONOMERS**

**@SO Analyt.Chem.; 68, No.22, 15th Nov.1996, p.3951-7**

**Résumé Document 00001305 ==> Ligne 11 Fac = 1.88 Qualité = 74.13 Contribution = 6.73**

**@TI STRUCTURE AND PROPERTIES OF THE COMPOSITE MEMBRANE OF REGENERATED SILK FIBROIN AND PVAL AND ITS APPLICATION TO AMPEROMETRIC TETRATHIAFULVALENE-MEDIATING GLUCOSE SENSOR**

**@SO J.Macromol.Sci.A; A33, No.2, 1996, p.209-19**

**Résumé Document 00001396 ==> Ligne 11 Fac = 1.81 Qualité = 72.21 Contribution = 7.50**

**@TI MACROMOLECULAR WIRING OF OXIDOREDUCTASES AND POTENTIAL INTERESTING APPLICATIONS**

**@SO Trends in Polymer Science; 3, No.10, Oct.1995, p.342-7**

**Résumé Document 00001402 ==> Ligne 13 Fac = 1.51 Qualité = 77.89 Contribution = 8.99**

**@TI PHOTOCURABLE POLYMER MATRICES FOR POTASSIUM-SENSITIVE ION-SELECTIVE ELECTRODE MEMBRANES**

**@SO Analyt.Chem.; 67, No.19, 1st Oct.1995, p.3589-95**

Par opposition à ce facteur évoquant les propriétés de certains polymères, le reste de l'analyse est consacrée aux capteurs dans les procédés de transformation. Ainsi, la partie négative de ce même facteur :

### Analyse 2-1 (suite)

**IL Y A 41 MOTS EXPLICATIFS :**

**POUR LE FACTEUR 1 LISTE DES MOTS A COORDONNEE NEGATIVE**

**SEUILS POUR LA CONTRIBUTION 2\*MOY = 2 08 3\*MOY = 3 12**

LIGNE	FACTEUR	QUALITE	CONTRI		
25	-0 40	58.67	5 83	676	PROCESS
25	-0 42	38.48	4 22	870	SYSTEM
25	-0 45	36.64	5 30	878	TEMPERATURE
26	-0 61	114.79	15 05	157	CONTROL
26	-0 63	31.00	4 00	302	EXTRUSION
26	-0 53	20.87	2 66	514	MELT
26	-0 49	19.68	2 17	706	QUALITY
27	-0 79	11.78	2 14	56	AUTOMATION
27	-0 76	28.56	3 22	92	CAVITY
27	-0 74	91.57	12 29	427	INJECTION
27	-0 78	63.55	9 31	546	MOULD
27	-0 64	97.30	11 55	548	MOULDING
27	-0 77	23.07	3 24	626	PLASTIC
27	-0 79	104.70	15 72	667	PRESSURE
28	-0 95	21.06	2 52	73	BLOW
28	-0 92	29.86	2 49	166	COOLING
28	-0 95	42.54	5 75	225	DEVICE
28	-0 86	54.58	6 95	483	MACHINE
28	-0 86	20.19	2 20	485	MACHINES
28	-0 92	24.38	3 15	924	UNIT
29	-0 95	24.55	2 56	381	HEATING
29	-1 05	31.83	2 76	649	POSITION
29	-0 96	38.08	2 33	825	SO
30	-1 22	16.49	2 24	13	ACTUATOR
30	-1 22	28.01	3 24	394	HYDRAULIC
30	-1 17	64.04	4 35	410	INCLUDES
31	-1 35	33.33	3 11	48	ASSEMBLY
32	-1 48	42.11	2 14	658	PREDETERMINED
33	-1 66	27.90	2 90	552	MOVABLE
33	-1 63	23.79	2 76	600	PAIR
33	-1 69	39.88	4 70	903	TOOL
33	-1 61	34.49	4 77	951	WASTE
34	-1 78	15.22	3 52	630	PLATEN
35	-1 91	30.13	4 55	170	CORE
35	-1 92	36.08	4 57	391	HOUSING
35	-1 99	44.58	7 23	709	RAM
35	-1 90	37.01	3 78	950	WALLS
36	-2 06	51.45	9 98	94	CHAMBER
37	-2 26	56.91	5 00	933	VALVE
44	-3 37	62.19	16 35	763	ROD
50	-4 34	55.36	18 49	568	NUT

## CAPTEURS Mise à jour 2-2

*Même constat pour la version 2 de l'analyse capteurs mise à jour*

**IL Y A 35 MOTS EXPLICATIFS :**

**POUR LE FACTEUR 1 LISTE DES MOTS A COORDONNEE POSITIVE** [*polymères utilisés comme biocapteurs*]

**SEUILS POUR LA CONTRIBUTION 2\*MOY = 5 29 3\*MOY = 7 94**

LIGNE	FACTEUR	QUALITE	CONTRI		
1	3 69	149 69	28 27	<b>228</b>	<b>OXIDASE</b>
3	3 45	116 96	19 93	<b>9</b>	<b>AMPEROMETRIC</b>
6	3 20	145 29	43 49	<b>131</b>	<b>GLUCOSE</b>
8	2 94	58 94	13 68	<b>21</b>	<b>BIOSENSORS</b>
8	2.99	114 13	25 54	<b>100</b>	<b>ENZYME</b>
8	2 93	112 18	18 60	<b>155</b>	<b>IMMOBILISATION</b>
9	2 79	158.06	27 58	<b>94</b>	<b>ELECTRODE</b>
10	2 70	38 10	8.62	<b>258</b>	<b>POLYTHIOPHENE</b>
11	2 61	58.97	12 75	<b>282</b>	<b>REDOX</b>
12	2 48	150 43	31 04	<b>195</b>	<b>MEMBRANE</b>
12	2 48	120 58	24 34	<b>196</b>	<b>MEMBRANES</b>
12	2 53	26 65	7 57	<b>259</b>	<b>POROUS</b>
13	2 42	21 76	9 24	<b>172</b>	<b>IONS</b>
13	2 42	93 56	26 66	<b>257</b>	<b>POLYPYRROLE</b>
14	2 28	108 57	16 48	<b>95</b>	<b>ELECTRODES</b>
15	2 18	45 10	7 04	<b>198</b>	<b>METHACRYLATE</b>
16	2.07	82 72	19 03	<b>171</b>	<b>ION</b>
16	2.08	54 38	13 26	<b>236</b>	<b>PH</b>
17	1 96	38 29	8 34	<b>1</b>	<b>ACID</b>
17	2 01	53 52	18 02	<b>229</b>	<b>OXYGEN</b>
17	2 03	28.00	7 73	<b>241</b>	<b>PLASMA</b>
17	1 96	41 82	10.97	<b>250</b>	<b>POLY</b>
17	2 01	91 12	16 78	<b>304</b>	<b>SELECTIVE</b>
18	1 91	35 37	6 83	<b>129</b>	<b>GEL</b>
18	1 90	28 58	6 78	<b>234</b>	<b>PERMEABILITY</b>
20	1 64	58 03	13 32	<b>49</b>	<b>CONDUCTING</b>
21	1 52	61 24	9 58	<b>226</b>	<b>ORGANIC</b>
21	1 53	41 27	8 29	<b>366</b>	<b>VAPOUR</b>
23	1 33	82 87	16 44	<b>116</b>	<b>FILMS</b>
24	1 26	18 13	5 94	<b>150</b>	<b>HUMIDITY</b>
24	1 23	78 78	17 17	<b>256</b>	<b>POLYMERS</b>
25	1 08	25 72	5 69	<b>2</b>	<b>ACOUSTIC</b>
25	1 10	58.43	9 63	<b>34</b>	<b>CHEMICAL</b>
25	1.14	34.43	6 33	<b>255</b>	<b>POLYMERISATION</b>
27	0 92	128 10	23 27	<b>254</b>	<b>POLYMER</b>

**QUALITE = 67 ( 27 ) doc**

*On retrouve l'opposition avec le thème des capteurs utilisés dans les procédés de transformation pour leur contrôle*

**IL Y A 8 MOTS EXPLICATIFS ·**

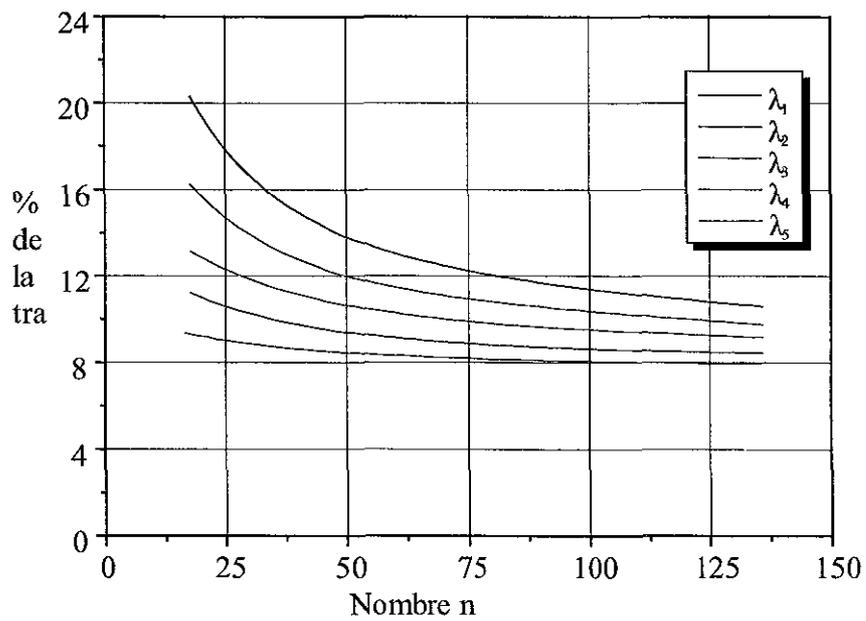
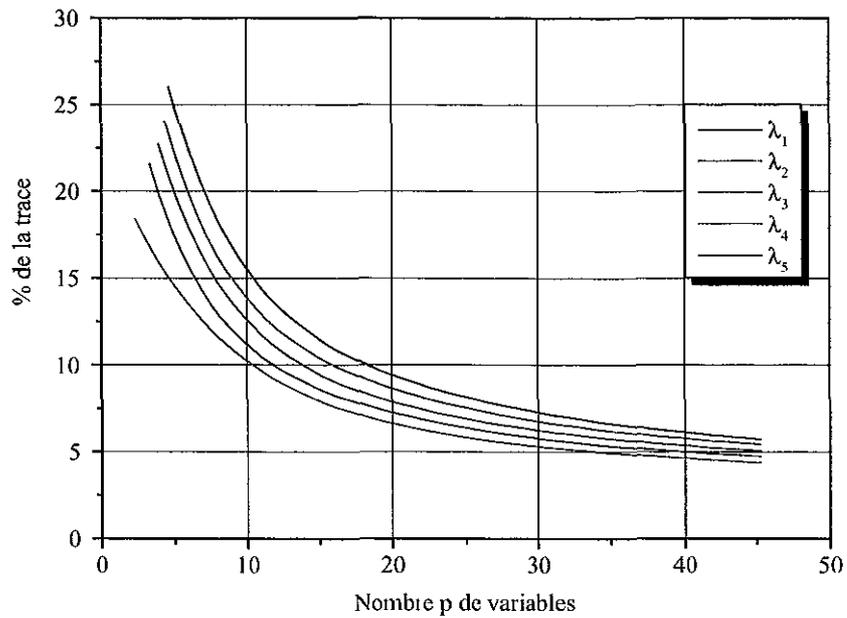
**POUR LE FACTEUR 1 LISTE DES MOTS A COORDONNEE NEGATIVE**

**SEUILS POUR LA CONTRIBUTION 2\*MOY = 5.29 3\*MOY = 7.94**

LIGNE	FACTEUR	QUALITE	CONTRI		
39	-0.39	32.02	5.30	<b>337</b>	<b>SYSTEM</b>
39	-0.41	33.47	6.65	<b>342</b>	<b>TEMPERATURE</b>
41	-0.57	104.47	19.62	<b>56</b>	<b>CONTROL</b>
42	-0.70	120.38	20.95	<b>213</b>	<b>MOULDING</b>
42	-0.73	99.33	20.23	<b>264</b>	<b>PRESSURE</b>
43	-0.74	102.63	18.51	<b>163</b>	<b>INJECTION</b>
43	-0.83	54.78	9.92	<b>184</b>	<b>MACHINE</b>
43	-0.75	61.40	12.87	<b>212</b>	<b>MOULD</b>

### ANNEXE III-A

Evolution des pourcentages d'inertie des cinq premières valeurs propres en fonction soit du nombre de variables, soit du nombre d'observations  
d'après Lebart ; Morineau ; Fénelon [1979]



**ANNEXE III-B**

A D D A D - 89 -

ANALYSE DES CORRESPONDANCES (AN CORR)  
D'APRES : YAGOLNITZER ET TABET

ANALYSE 296

INS 1 - TITRE :

TITRE ANALYSE DES MOTS ACTR > 1.34 (UNE FOIS LA MOYENNE)

LES VALEURS PROPRES VAL(1)= 1 00000

-----!NUM !VAL PROP!POURC!CUMUL!VARIAT!\*! HISTOGRAMME DES VALEURS PROPRES

NUM	VAL PROP	POURC	CUMUL	VARIAT	*
2	62452	883	883	0	000
3	50636	527	1410	0	356
4	40151	210	1620	0	316
5	38621	164	1784	0	046
6	37454	129	1913	0	035
7	36689	106	2019	0	023
8	36048	87	2106	0	019
9	35486	70	2176	0	017
10	34857	51	2227	0	019
11	33808	19	2246	0	032
12	33527	11	2257	0	008
13	32831	9	2266	0	021
14	32371	9	2275	0	014
15	32234	9	2284	0	004
16	31518	9	2293	0	022
17	31282	9	2302	0	007
18	31044	9	2311	0	007
19	30764	9	2320	0	008
20	30332	9	2329	0	013
21	29761	8	2338	0	017
22	29144	8	2347	0	019
23	28842	8	2356	0	009
24	28402	8	2365	0	013
25	27724	8	2374	0	020
26	27592	8	2383	0	004
27	27339	8	2392	0	008
28	27018	8	2401	0	010
29	26955	8	2410	0	002
30	26451	7	2419	0	015
31	26290	7	2428	0	005
32	26062	7	2437	0	007
33	25790	7	2446	0	008
34	25635	7	2455	0	005
35	24954	7	2464	0	021
36	24675	7	2473	0	008
37	24237	7	2482	0	013
38	23939	7	2491	0	009
39	23766	7	2500	0	005
40	23495	7	2509	0	008
41	23339	7	2518	0	005
42	23213	7	2527	0	004
43	22915	6	2536	0	009
44	22676	6	2545	0	007
45	22355	6	2554	0	010
46	22107	6	2563	0	007
47	22043	6	2572	0	002
48	21637	6	2581	0	012
49	21381	6	2590	0	008
50	21342	6	2599	0	001
51	21250	6	2608	0	003
52	21005	6	2617	0	007
53	20732	6	2626	0	008
54	20575	6	2635	0	005
55	20497	6	2644	0	002
56	20387	6	2653	0	003
57	19840	5	2662	0	016
58	19771	5	2671	0	002
59	19560	5	2680	0	006
60	19432	5	2689	0	004

## ANNEXE°III-C

### ANALYSES EFFECTUEES DURANT LA THESE

**RECHERCHE PAR REQUETES LOGIQUES** sur l'ensemble des analyses acces

#### 1. PROJET CAPTEURS BASE "RAPRA" evolution temporelle

Analyses factorielles

- 1 evolution des 106 mots sur en 4 groupes d'annees 1980 1997

#### 2- MISE A JOUR 97

**RECHERCHE PAR REQUETES LOGIQUES** acces a la base M.A.J. et les facteurs

Analyses factorielles

**2-1 : 987 DOCUMENTS, 933 MOTS SELECTION** frequence a 1% des docs soit 11

- 2 fac 1 a 10
- 3 fac 11 a 20
- 4 fac 21 a 30

**2-2 : 987 DOCUMENTS, 378 MOTS SELECTION CTR 1\*M SUR ANALYSE DES 933 MOTS**

**EDITION DES MOTS DONT LA CTR 2\*M**

- 5 fac 1 a 10
- 6 fac 11 a 20
- 7 fac 21 a 30

**2-3 : 987 DOCUMENTS, 378 MOTS SELECTION CTR 1\*M SUR ANALYSE DES 933 MOTS**

**EDITION DES MOTS DONT LA CTR 3\*M**

- 8 fac 1 a 10
- 9 fac 11 a 20
- 10 fac 21 a 30

#### 3- PROJET CAPTEUR BASE RAPRA

Analyses factorielles

- 11 fac 1 a 10
12. fac 11 a 20

**Classifications automatiques** classification sur les mots des résumés

#### 4- PROJET CAPTEURS (BASE 2) Base ISI

### **Analyses factorielles**

- 13 fac 1 a 10
- 14 fac 11 a 20
- 15 fac 21 a 30

### **Classifications automatiques classification sur les mots des résumés**

## **5- COMPETITIVE INTELLIGENCE**

### **Analyses factorielles**

- 16. fac 1 a 10
- 17 fac 11 a 20
- 18 fac 21 a 30

### **Classifications automatiques classification sui les mots des résumés**

## **6- POLYMERES A CRISTAUX LIQUIDES (LCP) analyses nouvelles**

### **6.1 Analyses factorielles LCP PURS 548 DOCUMENTS sans doublons et faux amis**

#### **SELECTION DES MOTS DE L'ANCIENNES ANALYSE**

- 19. fac 1 a 10
- 20 fac 11 a 20
- 21 fac 21 a 30

### **6.2 Analyses factorielles LCP PURS 548 DOCUMENTS sans doublons et faux amis**

#### **SELECTION DES 746 MOTS ayant une fiequence supetieure a 5**

- 22 fac 1 a 10
- 23 fac 11 a 20
- 24. fac 21 a 30

### **6.3 Analyses factorielles LCP PURS 548 DOCUMENTS sans doublons et faux amis**

#### **SELECTION DES 296 MOTS ayant unr contribution superieure a 1 34**

- 25 fac 1 a 10
- 26 fac 11 a 20
- 27 fac 21 a 30

### **6.4 Analyses factorielles LCP PURS 548 DOCUMENTS sans doublons et faux amis**

#### **SELECTION DES 673 MOTS ayant une frequence d'OCCURENCE superieure a 5**

- 28. fac 1 a 10
- 29 fac 11 a 20
- 30 fac 21 a 30

### **6.6 Analyses factorielles LCP PURS LES 345 DOCUMENTS REJETES**

#### **489 MOTS de frequence a 5**

- 31. fac 1 a 10
- 32. fac 11 a 20
- 33. fac 21 a 30

### **6.7 Analyses factorielles LCP PURS LES 345 DOCUMENTS REJETES**

192 mots d'inertie superieure a 1\*Moyenne dans l'analyse precedente (des 489 MOTS de frequence a 5)

- 37 fac 1 a 10
- 38 fac 11 a 20
- 39. fac 21 a 30

### **6.8 Analyses factorielles LCP PURS LES 345 DOCUMENTS REJETES**

727 MOTS SUR LES 746 SONT PRESENTS DANS CETTE SELECTION

- 40. fac 1 a 10
- 41 fac 11 a 20
- 42 fac 21 a 30

## **7- LCP: POLYMERES A CRISTAUX LIQUIDES Anciennes analyses**

**RECHERCHE PAR REQUETES LOGIQUES LCP PURS c'est-à-dire sans doublons et faux amis**

### **Analyses factorielles LCP PURS 508 DOCUMENTS**

**7-1 508 DOCUMENTS lcp pur, 769 MOTS SELECTION frequence a 1% des docs soit freq 5**

- 43. fac 1 a 10
- 44 fac 11 a 20
- 45. fac 21 a 30

**7-2 508 DOCUMENTS lcp pur, 671 MOTS SELECTION frequence d'occurence a 1% des docs soit freq 5**

- 46 fac 1 a 10
- 47 fac 11 a 20
- 48. fac 21 a 30

**7-3 508 DOCUMENTS, 358 MOTS SELECTION CTR 1\*M SUR ANALYSE DES 941 MOTS**

**EDITION DES MOTS DONT LA CTR 2\*M**

- 49 fac 1 a 10
- 50 fac 11 a 20
- 51. fac 21 a 30

## **8- JOURNAL BREF RHONE-ALPES**

**acces par requête documentaire acces**

**Analyses factorielles**

- 52. fac 1 a 10
- 53 fac 11 a 20
- 54 fac 21 a 30

## **9- RECYCLAGE DU POLYEURETHANE**

Accès par requête logique accés

**9-1 analyse 448 mots sélectionnés après dédoublement (528 doc) et analyse des 448 mots ayant une ctr 1 pour l'analyse des mots de fréquence 5 non représentés)**

55. fac 1 a 10

56. fac 11 a 20

57. fac 21 a 30

**9-2 analyse 261 mots sélectionnés de dimension à 8 d'après (528 doc.) et analyse des 448 mots**

58. fac 1 a 10

59. fac 11 a 20

60. fac 21 a 30

## **BIBLIOGRAPHIE**

## BIBLIOGRAPHIE

**Aaker D. A [1983]**, "Organizing a strategic information scanning system", *California Management Review*, vol. XXV, n°2, pp. 76-83

**Agosti M. ; Smeaton A.F. [1996]**, *Information retrieval and hypertext*, Kluwer Academic Publishers.

**Aguilar F.J. [1967]**, *Scanning the Business Environment*, Macmillan, New York.

**Alberto T. ; Combemale P. [1993]**, *Comprendre l'entreprise : théorie, gestion, relations sociales*, Nathan, Paris

**Ansoff I.H. [1965]**, *Corporate Strategy*, McGraw-Hill, New York.

**Ansoff I.H. [1980]**, "Strategic issue management", *Strategic Management Journal*, vol.1, pp. 131-148

**Bardin L. [1989]**, *L'Analyse de contenu*, PUF, Paris.

**Barré [1991]**, "Clustering Research Fields for Macro-strategic Analysis : a Comparative Specialisation approach", *Scientometrics*, Vol 22, n°1, p.95-112.

**Bartoli J.A. [1996] ; Le Moigne J.L. [1996]**, *Organisation intelligente et système d'information stratégique*, Economica, Paris.

**Baumard P. [1991]**, *Stratégie et surveillance des environnements concurrentiels*, Masson, Paris.

**Bell D. [1973]**, *The Coming of post-industrial society : a venture in social forecasting*, Basic books, New York

**Benzécri J.P [1973]**, *L'analyse des données*. Dunod, Paris : *La Taxinomie*, Vol.1. ; *L'analyse des données. L'Analyse des correspondances*, Vol.2.

**Benzécri J.P et alii [1980]** : *Pratique de l'analyse des données. Linguistique et Lexicologie*, Dunod, Paris, T. 3.

**Benzécri J.P [1984]** "Description des textes et analyse documentaire", *Cahiers de l'Analyse des Données*, n°2.

**Berelson B. ; Lazarsfeld P.F. [1948]**, *The Analysis of communications content*, University of Chicago et Columbia University, New York

**Berry M.W. ; Dumais S.T. [1994]**, "Using linear algebra for intelligent information retrieval", sur <http://superbook.bellcore.com/~std/lsi.papers>

**Besson B. ; Poussin J.C. [1996]**, *Du renseignement à l'intelligence économique. Détecter les menaces et les opportunités de l'entreprise*, Dunod, Paris.

**Bloch A. [1996]**, *L'Intelligence économique*, Economica, Paris

**Bouché R. ; Germain N. [1991]**, "Bibliométrie, infométrie et analyse automatique des documents écrits", *Revue Française de Bibliométrie*, n°9

**Bouché R. [1989]**, "Le syntagme nominal, une nouvelle approche des bases de données textuelles" in *Meta*, 1989

**Callon M. ; Laredo P. ; Mustar P. [1995]**, *La Gestion stratégique de la recherche et de la technologie. L'évaluation des programmes*, Economica, Paris.

**Callon M. , Leydesdorff L.[1987]**, "La recherche française est-elle en bonne santé?" *La Recherche* n°186, p.412-419.

**Carter C.F. ; Williams B.R. [1957]**, *Industry and technical progress*, Oxford University Press, London

**Charue-Duboc F. (sous la direction de) [1995]** : *Des Savoirs en Action. Contribution à la recherche en gestion*, Editions L'Harmattan.

**Chute C.G. ; Yang Y. [1995]**, "An Overview of statistical methods for the classification and retrieval of patient events", *Methods of information in Medicine*, n°34, pp. 104-110

**Cleland D.I. ; King W.R. [1975]**, "Competitive business intelligence systems", *Business Horizons*, Décembre 1975, pp. 19-28.

**Clerc P. [1997]**, "Intelligence économique : enjeux et perspectives", in *Rapport mondial sur l'information*, chapitre 22, sur <http://www.unesco.org.webworld/wirerpt>

**Confland D. [1997]**, *Economie de l'information spécialisée. Valeurs, usages professionnels, marchés.*, ADBS EDITIONS, Paris.

**Courtial J.P. [1990]**, *Introduction à la scientométrie. De la bibliométrie à la veille technologique*, Anthropos-Economica, Paris.

**Crozier M. [1963]**, *Le Phénomène bureaucratique*, Seuil, Paris.

**Deerwester S. et alii [1990]**, "Indexing by latent semantic analysis", *Journal of the American Society for Information Science [JASIS]*, vol. 41, n°6, pp. 391-407

**Deerwester S. [1988]**, "Improving information retrieval with latest semantic indexing", in *Information Technology : planning for the next fifty years*, Proceedings. ASIS 88, vol. 25, Atlanta (Georgia).

**Doré J.C. et alii [1996]**, "Correspondence factor analysis of the publications patterns of 48 countries over the period 1981-1992", *Journal of the American Society for Information Science [JASIS]*, vol. 47, n°8, pp. 588-602

**Desjardins-Bourcier et alii [1990]**, *Veille technologique, revue de la littérature et étude de terrain*, CERSI Villeurbanne Décembre 1990.

**Drèze X. [1995]**, "Utiliser des bases de données à des fins stratégiques", *Revue Française de Gestion*, Septembre-Octobre N°105 pp.76-84.

**Drucker P. [1956]**, "Managing the educated", in D.H.Fenn (ed.) *Management Missions in a new society*, McGraw-Hill, New York

**Ducasse R. [1978]**, *Méthodes du traitement des données bibliométriques pour la gestion des systèmes d'information. Application à l'analyse prévisionnelle de la demande d'ouvrages en bibliothèque*, Thèse en Sciences de l'Information et de la Communication, Université de Bordeaux 3.

**Dumais S.T. et alii [1988]**, "Using latent semantic analysis to improve access to textual information", in *Proceedings of Conference on Human Factors in Computing*, New York, ACM, pp. 281-285

**Dumais S.T. [1993]**, LSI meets TREC : a status report, Text Retrieval Conference [TREC], 1, March 1993, National Institute of Standards and Technology Special Publication.

**Durand J.P. [1994]**, "La Production sociale des systèmes d'information", *Sciences de la Société*, n°33, Octobre 1994, pp. 25-32.

**Eckart ; Young [1936]**, "The approximation of one matrix by another of lower rank" *Psychometrika*, 1, pp. 211-218

**Escofier [1988]**, *Analyses factorielles simples et multiples. Objectifs, méthodes et interprétation*, Dunod Bordas, Paris

**Estivals R. [1969]**, "La Statistique bibliographique", *Bulletin des bibliothèques de France*, n°12, pp. 481-502

**Favier L. [1994]**, *Bibliométrie et linguistique. Réflexions sur le couplage d'outils statistiques et de traitement automatique de la langue dans l'analyse du contenu de bases de données textuelles*, Mémoire de DEA, Enssib, Villeurbanne.

**Fénelon J.P. [1981]**, *Qu'est-ce que l'analyse des données ?*, Lefonen, Paris.

**Fievet Général G. [1992]**, *De la stratégie militaire à la stratégie d'entreprise*, Intereditions Paris.

**Fuld L.M. [1995]**, *The New competitor intelligence. The complete resource for finding, analysing, and using information about your competitors*, Wiley & Sons, New York.

**Gibbons P.T. ; Prescott J.E. [1996]**, "Parallel competitive intelligence processes in organisations", *International Journal Technology Management*, Vol. 11, n°1-2, pp. 162-178

**Godet M. [1997]** "La prospective stratégique. Prospective-stratégie: différences et complémentarités". *Futuribles* Avril 1997- N°219 p.45-59

**Golovchinsky G. [1993]**, *Queries-R-links : Query-based browsing in a full-text retrieval system*, Master's thesis, University of Toronto, Canada, sur <http://anarch.ie.utoronto.ca/people/golovch/publications/MASc/thesis.html>

**Gross P.L.K. ; Gross E.M. [1927]**, "College libraries and chemical education", *Science*, vol. 66, pp. 1229-1234

**Habert B. ; Nazarenko A. ; Salem A. [1997]**, *Les Linguistiques de corpus*,. A. Colin/Masson, Paris, 1997

**Hamlat A.A [1984]** "Analyse des répétitions et indexation automatique" Cahier de l'Analyse des Données, n°2.

**Harbulot C. [1992]**, *La Machine de guerre économique. Etats-Unis, Japon, Europe*, Economica, Paris.

**Hatem F. [1996]** Introduction à la Prospective, Editions Economica, Paris 112 p.

**Hassid L. ; Jacques-Gustave P. ; Moinet N. [1997]**, *Les PME face au défi de l'intelligence économique. Le renseignement sans complexe*, Dunod, Paris.

**Hendricks C. [1995]**, *Transferts de technologie inter-firmes: un renouvellement par l'analyse de la diffusion et de la création de la technologie*, Doctorat de Sciences économiques, Université de Nice, 1995

**Hume [1923]**, *Statistical bibliography in Relation to the Growth of Modern Civilisation*, Grafton

**JADT [1995]**, *3èmes Journées Internationales d'Analyse Statistique des Données Textuelles*, Rome 11-13 Décembre 1995, CISU, Roma, 2 volumes

**JADT 1998**, *4èmes Journées Internationales d'Analyses Statistiques des Données Textuelles* Nice, 1998 ; Editions UPRESA, Inalf, 98 bd E-Herriot, B.P. 209, F-06204 Nice cedex 3.

**Jakobiak F. [1991]**, *Pratiques de la veille technologique*, Editions d'Organisation, Paris.

**Laborit H. [1973]**, *Société informationnelle. Idées pour l'autogestion*, Ed. du Cerf, Paris.

**Lafouge T. [1993]**, "Bibliométrie : esquisse d'un panorama", *Sciences de la société*, n°29, pp. 99-110

**Lenoir R. et Prot B. [1979]** *L'information économique et sociale. Pour prévoir, produire et vivre en société. Tome II Rapport au Président de la République* La Documentation Française 327 pages.

**Lainé S. [1982]**, *Extraction et sélection de descripteurs complexes dans un ensemble de textes pour leur indexation automatique*, thèse, Université de Lyon 1.

**Lainée F. [1991]**, *La Veille technologique. De l'amateurisme au professionnalisme*, Eyrolles, Paris.

**Landauer T.K. ; Littman M. [1995]**, "A statistical method for language-independent representation of the topical content of text segments" sur <http://superbook.bellcore.com/~std/lis.papers>

**Lebart L. ; Salem A. [1994]**, *Statistique textuelle*, Dunod, Paris.

**Lebart L. ; Morineau A. ; Fenelon J.P. [1979]**, *Traitement des données statistiques: Méthodes et programmes*, Dunod, Paris.

**Lesca H. [1986]**, *Système d'information pour le management stratégique de l'entreprise. L'entreprise intelligente*, McGraw-Hill, Paris.

**Lesca H. [1992]** "Pour un management stratégique de l'information", *Revue Française de Gestion*, Septembre-Octobre 1992, pp.54-63.

**Lesca H. [1994]**, *Veille stratégique. L'intelligence de l'entreprise*, Aster, BP 2131 69003 Villeurbanne cedex.

**Lochbaum K.E. ; Streeter L.A. [1989]**, "Comparing and combining the effectiveness of latent semantic indexing and the ordinary vector space model for information retrieval", *Information Processing & Management*, vol. 25, n°6, pp. : 665-676

**Livre vert [1996]**, *Livre vert sur l'innovation*, Commission Européenne.

**Marchesnay M. [1993]**, *Management stratégique*, Eyrolles, Paris.

**Martinet B. ; Marti Y.M. [1995]**, *L'Intelligence économique. Les yeux et les oreilles de l'entreprise*, Editions d'Organisation, Paris.

**Martinet B. ; Ribault J.M. [1989]**, *La Veille technologique, concurrentielle et commerciale. Sources, Méthodologie, Organisation*, Les Editions d'Organisation

**Mayère A. [1993]**, "Sciences de gestion et sciences de l'information : fragments d'un discours inachevé", *Revue Française de Gestion*, Novembre-Décembre 1993, pp. 102-116

**Mayer R. [1990]** *Information et compétitivité*. Rapport du groupe présidé par R.Mayer. La France, L'Europe. Xème Plan 1989-1992. Ed : La Documentation Française 302 pages.

**Messina R. ; Eliot J.Y [1995]**, "L'Ouverture et la chasse : réflexion sur l'intelligence économique", *Technologies Internationales*, n°18, Octobre 1995, pp. 41-44

**Miller R.E. [1971]** *Innovation, Organisation and Environment*, Sherbrooke, Canada: University of Sherbrooke

**Mintzberg H. [1994]**, *Grandeur et décadence de la planification stratégique*, Dunod, Paris.

**Mira S. 1993** "Le Rôle effectif du système d'information dans l'entreprise industrielle", *Revue Française de Gestion*, Septembre-Octobre 1993, pp. 36-43.

**Morin J. [1985]**, *L'Excellence technologique*, Publi-Union/J. Picollec, Paris

**Morse P.M. [1968]**, *Library effectiveness*, MIT Press, Cambridge.

**Nora S, Minc A. 1978**, L'informatisation de la société, Paris La Documentation Française, 163 p., 4 vol.

**Norton B. 1994** "Organizational cultures and their relevance to the information organization", *Aslib Proceedings*, vol. 46, n°6, Juin 1994, pp. 173-176.

**Okubo Y. [1996]**, "Le japon mise sur la recherche de base", *La Recherche*, n°292, pp. 44-48.

**Olsen [1974] R.P.**, *Sources of Process Innovation in the Textile Industry*, DBA dissertation, Havard Business School

**Pearce F.T. [1976]**, "Business intelligence systems : the need, development, and integration", *Industrial Marketing Management*, 5, pp. 115-138.

**Plan [X<sup>ème</sup> plan. 1990]**, *Information et compétitivité*, La Documentation Française, Paris.

**Plan [XI<sup>ème</sup> plan. 1994]**, *Intelligence économique et stratégie des entreprises*, La Documentation Française, Paris.

**Porter M. [1982]** [1980 pour l'édition originale], *Choix stratégique et concurrence. Technique d'analyse des secteurs et de la concurrence dans l'industrie*, Economica, Paris.

**Pritchard A.L. [1969]**, "Statistical bibliography or bibliometrics ?", *Journal of Documentation*, n°25.

**Quoniam L.[1992]**, "Bibliométrie sur des références bibliographiques : méthodologie", in H. Desvals, H. Dou (ed.), *La veille technologique*, Dunod, Paris, pp. 244-252

**Régnier F. [1995]**, "Connaissances tacites: un rôle stratégique dans l'entreprise", *Revue Française de Gestion*, Septembre-Octobre , N°105, pp.127-133.

**Reinert M. [1987]**, "Classification descendante hiérarchique et analyse lexicale par contexte : application au corpus des poésies d'A. Rimbaud", *Bulletin de méthodologie sociologique*, n°13, pp. 53-90

**Reinert M. [1993]**, "Les "Mondes lexicaux" et leur "logique" à travers l'analyse statistique d'un corpus de récits de cauchemars", *Langage et société*, n°66, pp. 5-39

**Revue Française de gestion [1995]**, "Les Chemins du savoir de l'entreprise", n° spécial, Septembre-Octobre.

**Riller H. [1995]**, *La Fonction documentation des grandes entreprises françaises*, Rapport de synthèse obtenu auprès de H.Riller 3 rue de l'Arrivée, 75749 Paris cedex 15.

**Rostaing H. [1996]**, *La Bibliométrie et ses techniques*, Sciences de la Société, Toulouse/Centre de recherche rétrospective de Marseille.

**Saglio J. [1991]**, "Echange social et identité collective dans les systèmes industriels", *Sociologie du travail*, n°4, pp. 529-544.

**Salmon R. ; Lineares Y. de [1997]**, *L'Intelligence compétitive. Une combinaison subtile pour gagner ensemble*, Economica, Paris.

**Salmon R.[1967]** *L'information économique, clé de la prospérité: un inventaire critique des moyens actuels d'information économique en France*. Editions Hachette, Collection Entreprise, Paris, 199 pages.

**Scott Morton M.S. [1995]**, *L'Entreprise compétitive au futur : technologie de l'information et transformation de l'organisation*, Edition d'Organisation, Paris.

**Simon H. [1990]**, *Le Nouveau management : la décision par les ordinateurs*, Economica, Paris.

**Spark-Jones K. [1972]**, "A Statistical interpretation of term specificity and its applications and retrieval", *Journal of Documentation*, pp. 11-21.

**Thévenot J. ; France-Lanord B. [1993]** "Système d'information : un précurseur", *Revue Française de Gestion*, Juin-Juillet-Août 1992, pp. 96-111.

**Villain J. [1990]**, *L'Entreprise aux aguets. Information, surveillance de l'environnement, propriété et protection industrielles, espionnage et contre-espionnage au service de la compétitivité*, Masson, Paris.

**Werner E. [1996]**, "Entreprises et information : le discours à l'épreuve des faits", *Technologies internationales*, n°22, Février 1996, pp. **XXX**

**Wiseman C. [1987]**, *L'informatique stratégique, nouvel atout de la compétitivité*, Ed. d'Organisation, Paris.

**Wolek F.W. [1977]**, "Uses and benefits of technical information systems", *Research Management*, Septembre 1977, pp. 37-41.

**Yolin J.M. [1998]**, *Internet et les PMI*, sur <http://www.ensmp.fr/industrie/jmycs>

## **LISTE DES TABLEAUX ET DES FIGURES**

## LISTE DES TABLEAUX ET DES FIGURES

<u>Numérotation et intitulé des tableaux</u>	<u>Page</u>
<b>Tab. 1</b> Nombre d'occurrences de "competitive intelligence"	22
<b>Tab. 2</b> Date et source des premières occurrences des termes "competitive intelligence", "business intelligence", "environmental scanning"	23
<b>Tab. 3</b> Date et source des premières occurrences des termes "intelligence économique", "veille technologique", "veille stratégique" dans la base DELPHES	24
<b>Tab. 4</b> Date charnière d'évolution des termes pour les bases issues de la catégorie MULTIIND	25
<b>Tab. 5</b> Nombre d'occurrences des termes "competitive intelligence", "business intelligence", "environmental scanning" dans les bases de données en sciences sociales	27
<b>Tab. 6</b> "Some of the many sources for strategic information"	29
<b>Tab. 7</b> Distinction entre les termes "data", "information", "intelligence"	36
<b>Tab. 8</b> Comparaison de la notion "d'intelligence économique" en France et aux Etats-Unis	43
<b>Tab. 9</b> Le modèle SOWT	47
<b>Tab. 10</b> La matrice du BCG	48
<b>Tab. 11</b> Deuxième version de la matrice du BCG	49
<b>Tab. 12</b> Les cinq forces qui déterminent la rentabilité de l'industrie	52
<b>Tab. 13</b> Les cinq paramètres qui commandent la concurrence au sein d'un secteur et les quatre types de veille associés	53
<b>Tab. 14</b> "Les fonctions d'un système de renseignement sur la concurrence"	54
<b>Tab. 15</b> La veille technologique selon Morin	64
<b>Tab. 16</b> Analyse Factorielle des Correspondances : Intelligence économique. Carte des facteurs 2 et 3.	67
<b>Tab. 17</b> Besoin d'information / Compétences requises	84
<b>Tab. 18</b> Loi de Zipf	89
<b>Tab. 19</b> Equivalence distributionnelle	105
<b>Tab. 20</b> Equivalence distributionnelle suite	105
<b>Tab. 21</b> Illustration de la déficience des méthodes documentaires selon les auteurs de la LSA	114
<b>Tab. 22</b> Tableau des données traitées par Berry et Dumais	116

<u>Numérotation et intitulé des figures</u>	<u>Pages</u>
<b>Fig 1</b> Coeur et dispersion d'une distribution bibliométrique	88
<b>Fig 2</b> Graphe de co-citations	98
<b>Fig 3</b> Diagramme stratégique	99
<b>Fig 4</b> Analyse par SVD d'après Berry et Dumais, et position de la requête "application theory"	119
<b>Fig 5</b> Analyse par AFC du même tableau de données que celui de Berry et Dumais, mais sans positionner la requête	120
<b>Fig 6</b> Position de la requête (point QUERY) "application theory" avec le traitement des données par AFC	121
<b>Fig 7</b> Analyse par SVD du même tableau en ajoutant les documents B18 B19 B20 au même titre que les autres documents	122
<b>Fig 8</b> Analyse par SVD d'après Berry et Dumais en ajoutant les documents B18 B19 B20 "using the SVD of a reconstructed term-document matrix"	123
<b>Fig 9</b> AFC sur les données de Berry et Dumais en ajoutant les documents B18 B19 B20 au même titre que les autres documents	124
<b>Fig 10</b> AFC sur les données de Berry et Dumais en ajoutant les documents B18 B19 B20 en tant qu'éléments supplémentaires	125
<b>Fig 11</b> Proposition de système documentaire	179