L'exploitation systématique des bases de données: des analyses stratégiques pour l'entreprise

ROSTAING Hervé, NIVOL William, QUONIAM Luc (1)
BEDECARRAX Chantal, HUOT Charles (2)

(1) Centre de Recherche Rétrospective de Marseille
Aix-Marseille III, Faculté St Jérome, 13397 Marseille CEDEX 13

(2) Centre Européen Scientifique en Mathématiques Appliquées
IBM, 68/76 quai de la Rapée, 75592 Paris CEDEX 12

INTRODUCTION

Aujourd'hui tout industriel se pose des questions sur l'environnement scientifique, technologique et concurrentiel de son entreprise. Pour répondre à ces questions il met en place une structure dite de veille technologique ou de veille stratégique. Cette structure doit faire preuve d'une vigilance de tous les instants et être à même de répondre à des questions assez générales sur l'état de la concurrence dans un domaine, d'établir des cartographies d'une technologie ou encore d'analyser des évolutions à travers le temps. Le système de surveillance à instaurer dans une entreprise doit donc assurer deux fonctions (*illustration 1*):

- contrôle en continu de l'environnement pour alerter en cas de menaces émergentes
- monter des dossiers thématiques pour connaître les tendances d'évolution et se positionner par rapport à ses concurrents

La première fonction demande essentiellement un contrôle de l'information informelle et floue tandis que la seconde impose une maîtrise de l'information formalisée et confirmée. La bibliométrie est un outil qui peut s'avérer très utile dans cette seconde tâche [1]. Son application va permettre l'élaboration d'indicateurs de tendances à partir du maximum de connaissances scientifiques ou technologiques que l'on puisse avoir sur un thème à un moment précis.

La bibliométrie est de plus en plus connue comme une méthode d'analyse des bases de données accessibles en ligne. La raison de ce recours aux bases de données est simple; elles représentent la première et la plus importante source d'information scientifique, technique et technologique [2]; elles offrent de plus l'avantage d'être analysables par des systèmes informatiques automatisés [3].

VEILLE TECHNOLOGIQUE - BIBLIOMETRIE

LE BESOIN INDUSTRIEL

CONNAISSANCE DE SON ENVIRONNEMENT
(technologique, scientifique, concurrentiel,...)

SYSTEMES DE SURVEILLANCE

- **EN CONTINU : ALERTER**
- **THEMATIQUES : POSITIONNEMENT / CONCURRENCE**

OUTIL : LA BIBLIOMETRIE

- **ANALYSE DE REFERENCES :**
  - *** INDICATEURS SYNTHETIQUES**

- **MASSE ETCOMPLEXITE DES DONNEES :**
  - *** OUTILS RAPIDES, SYSTEMATIQUES, APPROPRIES :**
    - **. SPECIFIQUES (Préparation des données)**
    - **. STATISTIQUES (Détermination d'une grille de lecture)**

- **INFORMATION CIBLEE = INFORMATION BREVETS**
  - *** NECESSAIRE AUX ENTREPRISES**
  - *** FIABILITE**

Nous allons présenter dans cet article une étude mettre en oeuvre deux outils adaptés au traitement automatique des données bibliographiques (*illustration 2*):

- Le premier automatise la manipulation des références pour restituer les caractéristiques bibliométriques du corpus étudié (distributions bibliométriques, listes de fréquences de termes, listes de fréquences de cooccurrences de termes, matrices d'occurrences, matrices de cooccurrences, matrices d'associations):

  DATAVIEW (développé au CRRM).

- Le second est l'outil statistique que l'on appliquera à ces caractéristiques bibliométriques pour discriminer et structurer le contenu informationnel du corpus:

  L'ANALYSE RELATIONNELLE (développée au CESMAP)

L'information concernant l'innovation industrielle est une information stratégique, aux yeux des entreprises. Il est crucial de pouvoir la maîtriser pour conserver sa performance dans l'art d'innover. L'information brevet en fait partie. Elle est la source de multiples renseignements sur l'état des innovations. Nous avons donc choisi de cibler cette étude sur l'information brevet à travers la base Derwent (WPIL).

Le thème choisi pour cet travail est une technologie charnière entre la médecine et la pharmacie: *Les systèmes transdermiques thérapeutiques sous forme de patchs (T.T.S).* On peut facilement envisager qu'un tel sujet puisse être sensible pour une entreprise pharmaceutique. Dans un système de veille, on pourrait imaginer que ce thème soit sélectionné comme étant l'un des facteurs critiques, c'est-à-dire un sujet stratégique pour garantir la pérennité de l'entreprise.

● OBJECTIF DE L'ETUDE:

Une évaluation méthodologique de nouveaux traitements bibliométriques

Une référence signalétique dans une base de données contient diverses informations réparties en plusieurs rubriques. Ces rubriques, nommées champs, ont des portées significatives différentes. Il est rare que la richesse de cette diversité d'information soit pleinement exploitée dans les études bibliométriques de corpus de références.

Traditionnellement, en bibliométrie, les analyses statistiques ne traitent que le contenu d'un champ à la fois. Les méthodes d'analyse de co-citations [4] établissent leurs cartes sur la rubrique concernant les citations faites par les auteurs scientifiques. La méthode de mots-associés [5] est développée pour exploiter un champ indexé. La méthode des citations-

LE CAS ETUDIE

LE CORPUS

- LE THEME

SYSTEMES TRANSDERMIQUES THERAPEUTIQUES SOUS FORME DE PATCH (T.T.S.)

SYSTEMES ADHESIFS QUI ASSURENT UNE DIFFUSION CONTROLEE D'UN PRINCIPE ACTIF PAR VOIE TRANSDERMIQUE

- LA SOURCE — LA BASE BREVETS DERWENT WPIL

LES TRAITEMENTS AUTOMATIQUES DES DONNEES TEXTUELLES

- **DATAVIEW (C.R.R.M.)**

LES TRAITEMENTS STATISTIQUES

- **L'ANALYSE RELATIONNELLE (C.E.M.A.P. - F.MARCOTORCHINO, P.MICHAUD)**

croisées de journaux [6] est une approche vers cette combinaison d'informations de champs différents. Cette dernière se base sur une matrice croisant des journaux "citants" (champ source) avec des journaux cités (champ citation).
Cette déficience dans les traitements bibliométriques est due à la complexité des relations engendrées par toutes les combinaisons de ces informations.

Cet article tente d'apporter une réponse à cette carence. Il traite de la complémentarité d'informations apportées par la prise en compte simultanée de trois champs de la base brevets de Derwent.
Tout brevet référencé dans la base est qualifié, au sens informationnel, par les codes de trois classifications différentes. Une classification documentaire découpe les domaines scientifiques en sections. Si la classification est construite à partir d'un principe de hiérarchie, ces sections sont alors elles-mêmes découpées en sous-sections, classes, sous-classes, groupes, sous-groupes... Chaque niveau dans cette hiérarchie de découpage est représenté par une codification.

Une référence de brevet dans la base Derwent reçoit donc l'affectation, dans plusieurs champs, de différents codes qui illustrent les thèmes abordés par l'invention.
Les trois "champs codes" que nous allons exploiter sont les champs:

- DC (Derwent Codes) : Classification documentaire établie par Derwent
- MC (Manuel Codes) : Autre classification documentaire établie par Derwent
- IC (International Patent Classification) : Classification établie par les instituts officiels de dépôts de brevets.

Pour estimer l'apport spécifique de chacune de ces classifications, nous allons les confronter pour le corpus de références brevets établi sur le thème des patchs transdermiques thérapeutiques. Les méthodes d'Analyse de Données Relationnelles font partie du panel des analyses statistiques élaborées dans le but de mieux cerner les phénomènes complexes. Leur exploitation va nous permettre d'évaluer deux caractéristiques des relations que ces champs peuvent entretenir (*illustration 3*):

- les **complémentarités** de ces classifications documentaires en qualité de descripteurs de brevets: nous allons estimer si le fait d'utiliser ces trois classifications simultanément nous permet de mieux décrire les réels liens entre les brevets.

- les **correspondances** ou les **similarités** entre les codes de ces classifications au niveau de leur sens: on pourrait ainsi connaître les recouvrements de signification, les codes synonymes et les complémentarités entre les classifications.

OBJECTIFS

LES BASES DE DONNEES BREVETS ACCESSIBLES EN LIGNE OFFRENT UNE DIVERSITE D'INFORMATION

→ INFORMATION

* STRUCTUREE
* RICHE
* DIFFICILE A CORRELER

A PARTIR D'UN OU PLUSIEURS CHAMPS DE DESCRIPTION

- ETUDE DE LA COMPLEMENTARITE DES DESCRIPTEURS

- ETUDE DE LA CORRESPONDANCE OU DE LA SIMILARITE DES DESCRIPTEURS

● CONSTITUTION DU CORPUS DES REFERENCES

Cette première étape, dans une analyse bibliométrique, est certainement celle qui influence le plus la validité des résultats. Cette validité est capitale lorsque les résultats rentrent dans un processus de décision.

Pour notre étude, l'objectif principal est de présenter une méthode permettant de prendre en compte la diversité du sens apporté par trois classifications utilisées sur les brevets. Néanmoins, il est indispensable, pour conclure sur la cohérence des résultats, de constituer un ensemble homogène de références tout en assurant une couverture acceptable du sujet.

Le corpus dégagé doit, autant que possible, être suffisamment large pour couvrir le thème étudié et suffisamment étroit pour présenter un "bruit" aussi faible que possible. Nous avons donc volontairement réduit les risques de bruits pour qu'ils ne viennent pas perturber l'interprétation des résultats statistiques. L'ensemble des références, qui a été collecté pour cette étude, n'est donc probablement pas exhaustif mais devrait posséder une propriété d'homogénéité.

Dans le cadre d'une étude de veille, cette étape serait obligatoirement conduite en présence d'experts du thème pour garantir la validité de la base des connaissances construite.

La stratégie d'interrogation a été affinée selon une méthode itérative de type "coups de sonde". Chaque itération permet, après lecture d'échantillons de références, de dégager de nouvelles pistes pour enrichir la stratégie. Cette itération est répétée tant que les échantillons, obtenus par les croisements des nouvelles pistes, laissent apparaître des références non pertinentes. Les échantillons ont été estimés pertinents pour la stratégie d'interrogation suivante:

QUESTION 1 :	PATCH ou PATCHES ou PATCHS	(2106)
QUESTION 2 :	1 et TRANSDERM:	(103)
QUESTION 3 :	1 et THERAPEUTIC:	(36)
QUESTION 4 :	1 et PERCUTANE:	(15)
QUESTION 5 :	1 et (DRUG ou DRUGS)	(102)
QUESTION 6 :	1 et MEDICIN:	(14)
QUESTION 7 :	2 ou 3 ou 4 ou 5 ou 6	(160)
QUESTION 8 :	7 sans (CARDIAC# à coté de PATCH##)	(159)
QUESTION 9 :	8 sans (VASCULAR# à coté de GRAFT#)	(158)
QUESTION 10:	9 sans (PATCH## à coté de GRAFT#)	(156)
QUESTION 11:	10 sans (FASTENING# à coté de PATCH##)	(155)
QUESTION 12:	11 sans (CARRY à coté de PATCH##)	(155)
QUESTION 13:	12 sans (CARRIES à coté de PATCH##)	(154)
QUESTION 14:	13 sans PROSTHES:	(148)
QUESTION 15:	14 sans CAMERA	(147)
QUESTION 16:	15 sans (TEST# à coté de PATCH##)	(146)

Remarque: les questions sont posées sur l'Index de Base de la base Derwent, c'est à dire sur les champs: Résumé, Résumé équivalent, Titre et Titre normalisé de Derwent. ({#} = Troncature courte. {:} = Troncature large)

L'examen de cette stratégie finale montre que la simple utilisation des termes "PATCH" et "THERAPEUTIC" ne nous permettait par de couvrir la totalité des brevets du domaine. Par contre l'emploi d'autres termes, pour élargir la recherche, laissait apparaître des références hors-sujet car le terme PATCH a des significations multiples et variées (rustine, greffon, pièce de prothèse, chute de film, test d'allergie...). Par conséquent, les brevets, qui détiennent pour ce terme un sens autre que celui recherché sont désélectionnés.
Cette stratégie d'interrogation permet de dégager un corpus de 146 brevets. Le télédéchargement de ces références est réalisé non seulement pour les trois champs des classifications mais aussi pour tous les champs qui fournissent des renseignements pouvant aider à la compréhension des résultats des analyses statistiques.

● TRAITEMENT DES CARACTERISTIQUES BIBLIOMETRIQUES DU CORPUS

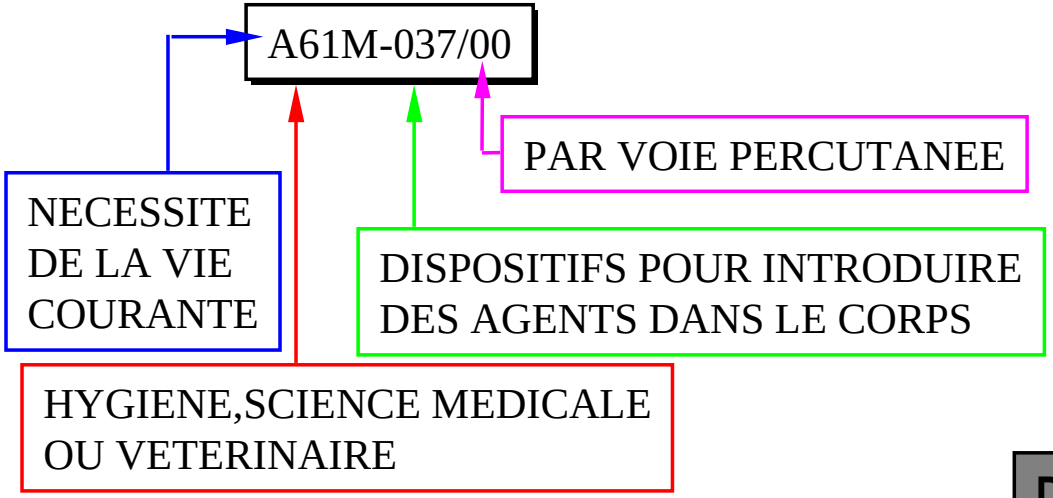
Pour confronter les différentes informations apportées par chacune des classifications nous allons reproduire leurs interactions par la construction de tableaux. Ces tableaux également appelés "matrices" sont le point d'entrée des analyses statistiques.

Choix des niveaux hiérarchiques:

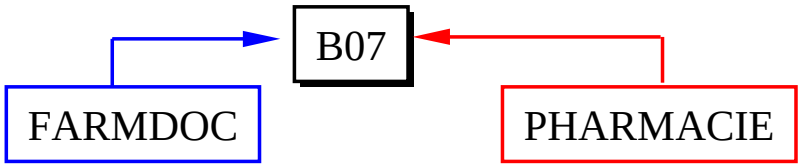
Les trois classifications brevets possèdent leur propre hiérarchie de codes. Un code est assimilable à un chemin pris parmi les branches de la hiérarchie. Dans cette hiérarchie, les branches sont réparties à partir d'un niveau de signification très large vers des niveaux de signification de plus en plus fins. Plus on descend dans les branches de la hiérarchie, plus le code a une représentation complexe et plus son sens est précis. Cette notion est représentée par l'*illustration 4* où l'on a expliqué la signification de chaque niveau de hiérarchie.

Quel niveau hiérarchique considérer pour cette étude?
Décider de prendre les niveaux hiérarchiques les plus précis pour avoir les descriptions les plus fines est contestable. En effet, tous les codes affectés aux références ne sont pas forcément renseignés jusqu'au dernier niveau de la hiérarchie. Par exemple sur notre corpus de références, 66 % des codes IPC sont renseignés jusqu'au dernier niveau contre 12 % pour les Manuels Codes.
Parallèlement, plus le niveau hiérarchique est fin, plus la diversité des codes augmente. Ceci est un argument en faveur du choix d'un niveau assez fin.

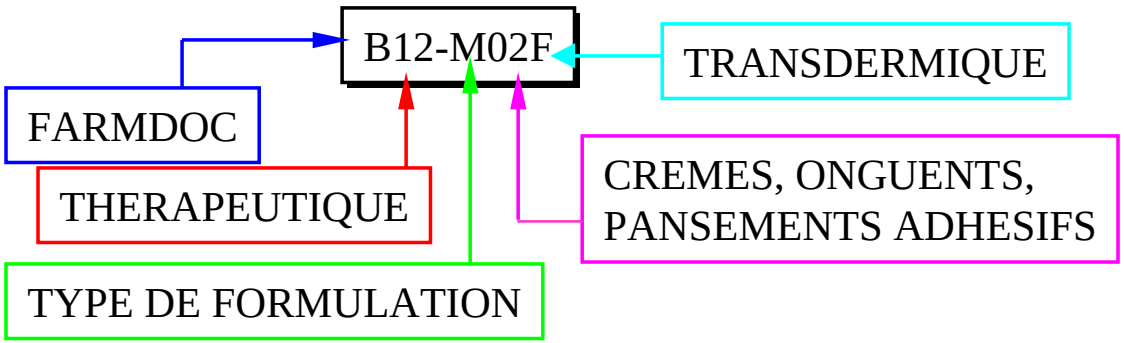
CODE C.I.B. CHAMP IC



DERWENT CODE CHAMP DC



MANUEL CODE CHAMP MC



Ce choix de niveaux de hiérarchie, impose donc un compromis entre la perte d'information et le gain de signification (pour un choix très fin, perte de certains codes mais codes restants plus précis).

Une pré-étude a été menée pour déterminer quels niveaux hiérarchiques satisfaisaient la meilleure solution statistique. Les critères étudiés pour chaque niveau hiérarchique étaient:

- le nombre de codes restants
- le nombre de brevets encore renseignés
- la qualité d'agrégation des brevets par l'Analyse Relationnelle pou ce niveau de la hiérarchie (nombre d'agrégats et nombre de codes non agrégés)

Remarque: Les codes à fréquence 1 ne sont pas considérés puisqu'ils n'établissent aucun lien entre les brevets.

Le choix s'est porté sur les niveaux hiérarchiques dont les nombres de codes et les nombres d'agrégats pour les différentes classifications documentaires sont proches (*illustration 5*). Les niveaux hiérarchiques choisis pour l'étude ont donc été:

- Les Derwent Codes à 3 caractères
- Les Manuel Codes à 3 caractères
- Les codes IPC à 7 caractères

L'absence, parmi ces critères purement statistiques, de critères qualitatifs pour comparer les degrés de sens à chaque niveau des hiérarchies est critiquable. En l'absence d'experts du domaine étudié, il était difficile de s'investir dans de telles considérations. Nous sommes conscients que ce choix demande à être confirmé par des critères plus qualitatifs mais nous voulions en premier lieu évaluer si cette utilisation "novatrice" de plusieurs codifications était bénéfique.

CHOIX DES NIVEAUX HIERARCHIQUES DES CODES

HIERARCHIES DES CODES (Nb de digits)	NB DE CODES	NB DE CODES APRES ELIMINATION DES CODES A FREQUENCE 1	NB DE BREVETS	NB DE BREVETS APRES ELIMINATION DES CODES A FREQUENCE 1	NB MOYEN DE CODES PAR BREVET	NB DE CLASSES	% DE CLASSES A UN ELEMENT
DC 3	52	32	146	146	3,6	27	30
IC 4	40	21	146	143	2,2	23	22
IC 7	94	35	146	140	2,9	42	36
IC 11	133	36	146	113	2	67	52
MC 3	51	42	143	143	4,6	41	34
MC 5	139	90	143	143	6,9	80	50
MC 7	315	158	142	141	8,7	106	74
MC8	144	69	135	134	4,3	80	55
MC 9	43	12	46	34	1,7	124	90

Construction des tableaux de l'étude

Le choix des niveaux étant fait pour chacune des hiérarchies, vient alors la phase de construction des tableaux pour l'analyse.

Plusieurs types de tableaux sont exploitables par les méthodes d'analyse statistique. En ce qui concerne l'Analyse Relationnelle des Données, les tableaux d'entrée sont du type matrice de présence-absence. Ces matrices croisent un ensemble *d'individus,* noté *I,* et un ensemble de *variables* descriptives, noté *J*. Les individus sont naturellement ici les 146 brevets et l'ensemble J des variables est constitué des codes. Le croisement de *I x J* fait donc figurer, à l'intersection de la ligne *i* et de la colonne *j,* la valeur 1 si le code *j* est présent dans la référence du brevet *i* et la valeur 0 dans le cas contraire. Ce tableau est la simple restitution des données élémentaires contenues dans les références rien n'est omis, ni ajouté, ni transformé.

Pour l'étude des complémentarités des codes, nous avons mené parallèlement l'analyse des regroupements des références:

- lorsqu'elles sont décrites par un des trois ensembles de codes
- lorsqu'elles sont décrites par la réunion des trois ensembles.

Nous avons donc construit quatre matrices; trois pour les ensembles isolés de codes et une par l'union de ces codes (*illustration 6*).

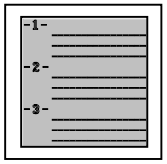
Ce passage des données textuelles (références) aux données tabulées (matrices) est réalisé par le logiciel DATAVIEW du CRRM.
Les matrices sont construites, en automatique, avec DATAVIEW selon la logique suivante:

- extraire tous les codes
- éliminer les codes qui ne sont pas renseignés jusqu'aux niveaux des hiérarchies choisies
- tronquer les codes restants à ces niveaux de hiérarchie
- construire les matrices de présence-absence

D'autres caractéristiques bibliométriques

L'*illustration 6 (à l'extrême droite)* présente aussi sous forme symbolique d'autres traitements bibliométriques que permet DATAVIEW. Ces résultats sous forme de distributions de fréquences ou de rangs de fréquences peuvent déjà apporter de nombreux renseignements. Il est possible d'examiner l'évolutions des brevets en fonction du temps, ce qui permet de situer le niveau de maturité du sujet étudié. De même, nous pouvons établir la répartition par pays de dépôts, ce qui peut permettre de dégager les pays dont les marchés sont convoités. On peut aussi connaître les sociétés leaders en nombre de dépôts de brevets. Ces résultats demandent souvent des prétraitements de reformatage, d'homogénéisation, de désambiguïsation

CONSTRUCTION ET ANALYSE DES MATRICES



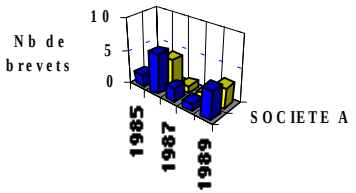
DATAVIEW

* DISTRIBUTION
* LISTES DE FREQUENCES

DC | MC 3 | IC 7 | DC +MC 3 +IC 7

Ref

REPARTITION PAR SOCIETE



REPARTITION PAR PAYS

ANALYSE RELATIONNELLE

MATRICES PERMUTEES

COMPLEMENTARITE DES CODES

DC +MC 3 +IC 7

Ref

DC | MC 3 | IC 7

Ref

CORRESPONDANCE DES CODES

DC +MC 3 +IC 7

Ref

difficilement envisageables sans outil informatique. DATAVIEW facilite ce travail préliminaire, souvent fastidieux, mais néanmoins fondamental pour garantir la qualité des données et donc la validité des résultats.

● L'ANALYSE STATISTIQUE

Pour évaluer la complémentarité des codes en qualité de description des brevets, l'Analyse Relationnelle est appliquée sur l'espace des brevets pour les regrouper par similarité de description de codes.
Pour évaluer les correspondances entre les codes, l'Analyse Relationnelle servira à regrouper les codes par leurs ressemblances de répartition dans l'ensemble des brevets.

La technique que nous allons mettre en oeuvre pour construire la partition des brevets en classes disjointes s'inscrit dans le cadre méthodologique général de l'Analyse Relationnelle. Nous avons pris l'option, dans cet article, de bannir toute formule mathématique et de réduire au minimum les explications et les justifications méthodologiques. On trouvera dans [1] et [7] l'essentiel de ces informations.
Nous nous contenterons de parler, ici, du point fondamental qui préside au déroulement du traitement, à savoir: le critère de classification qui mesure les similarités entre les objets que l'on compare. C'est incontestablement la question qu'il faut à tout prix se poser pour appliquer la procédure de classification dans un cadre clairement défini. Faute de quoi, on ne sait pas, a posteriori, expliquer la structure que l'on a mis en évidence.

L'agrégation des références brevets:

On va chercher ici a dégager des classes de brevets qui s'apparentent par les codes descripteurs qu'ils ont en commun, autrement dit des familles de brevets caractérisées par leurs similitudes en termes de technologies partagées.
Les spécialistes de la bibliométrie, de la scientométrie, de l'infométrie ou encore de la lexicographie mathématique, ont depuis fort longtemps mis en évidence et étudié les caractéristiques des distributions que l'on rencontre dans ces domaines (lois de Zipf, Bradford et Lotka).
Les données extraites des bases de données telles que celles dont nous disposons sont de cette nature. Ainsi, un certain nombre de codes se retrouvent dans la grande majorité des références (ce sont les thèmes qui ont présidé à la construction du corpus), d'autres apparaissent de façon moins systématique et d'autres enfin ne figurent que dans un faible, voire très faible, nombre de références. Il importe que le critère de classification qui va permettre de mesurer les ressemblances entre les brevets tienne compte de ces phénomènes.

Notre choix s'est porté sur le critère de Burt pondéré, qui outre ses bonnes propriétés axiomatiques, se base sur un indice de présence-rareté répondant parfaitement à notre problème. En effet, ce critère a pour effet d'attribuer aux codes un poids inversement proportionnel à leur présence dans le corpus. Ainsi, deux brevets qui partagent un code rare seront "plus similaire" que deux brevets qui auraient en commun un code présent dans de nombreuses références.

Les classes de brevets issues de la classification pilotée par le critère de Burt pondéré sont donc formées de brevets qui se ressemblent non seulement parce qu'ils partagent les mêmes codes mais encore parce que ces codes sont absents (ou peu "typiques") dans les autres brevets. On garantit le découpage du corpus en familles de brevets homogènes et discriminantes. L'analyse du résultat permet en outre d'expliquer chacune de ces classes en fonction des codes ou groupes de codes qui ont présidé à leur création, autrement dit, d'attacher à chacune d'elle une *étiquette* synthétique résumant les thèmes technologiques qu'elle recouvre.

La classification des codes

L'objectif de ce traitement est de découper l'ensemble des codes descripteurs en classes homogènes sur la base de leurs co-occurrences dans les références. Autrement dit, on cherche ici à dégager les pôles technologiques autour desquels s'articule le corpus.

Cette fois ce sont les colonnes du tableau de départ, c'est-à-dire l'ensemble des codes descripteurs, qui sont soumis au processus de classification.
Les considérations (lois Zipf-Bradeford-Lotka) qui avaient induit le choix du critère de classification sur les brevets sont encore valides dans ce contexte. Il n'est toutefois plus possible d'utiliser le critère de Burt pondéré. Celui-ci aurait en effet pour conséquence de faire jouer un rôle à la richesse de description des brevets. Or le fait qu'un brevet possède un plus ou moins grand nombre de codes ne doit pas être pénalisant. En revanche, il convient toujours de prendre en compte la fréquence d'apparition des codes qui, elle, reste tout à fait pertinente.
Notre choix s'est finalement porté sur le critère de Burt, très couramment utilisé en Analyse Relationnelle pour ses bonnes propriétés de règle d'agrégation.
A l'issue de ce traitement, nous obtenons une partition des codes descripteurs. Chacune des classes de cette partition regroupe un certain nombre de codes qui ont pour caractéristique d'apparaître conjointement dans les références de brevets.
On a donc mis en évidence des combinaisons de technologies qui présentent un fort taux de corrélation à l'échelle du corpus étudié.

● DISCUSSION DES RESULTATS

Etude de la complémentarité des codes pour décrire les brevets

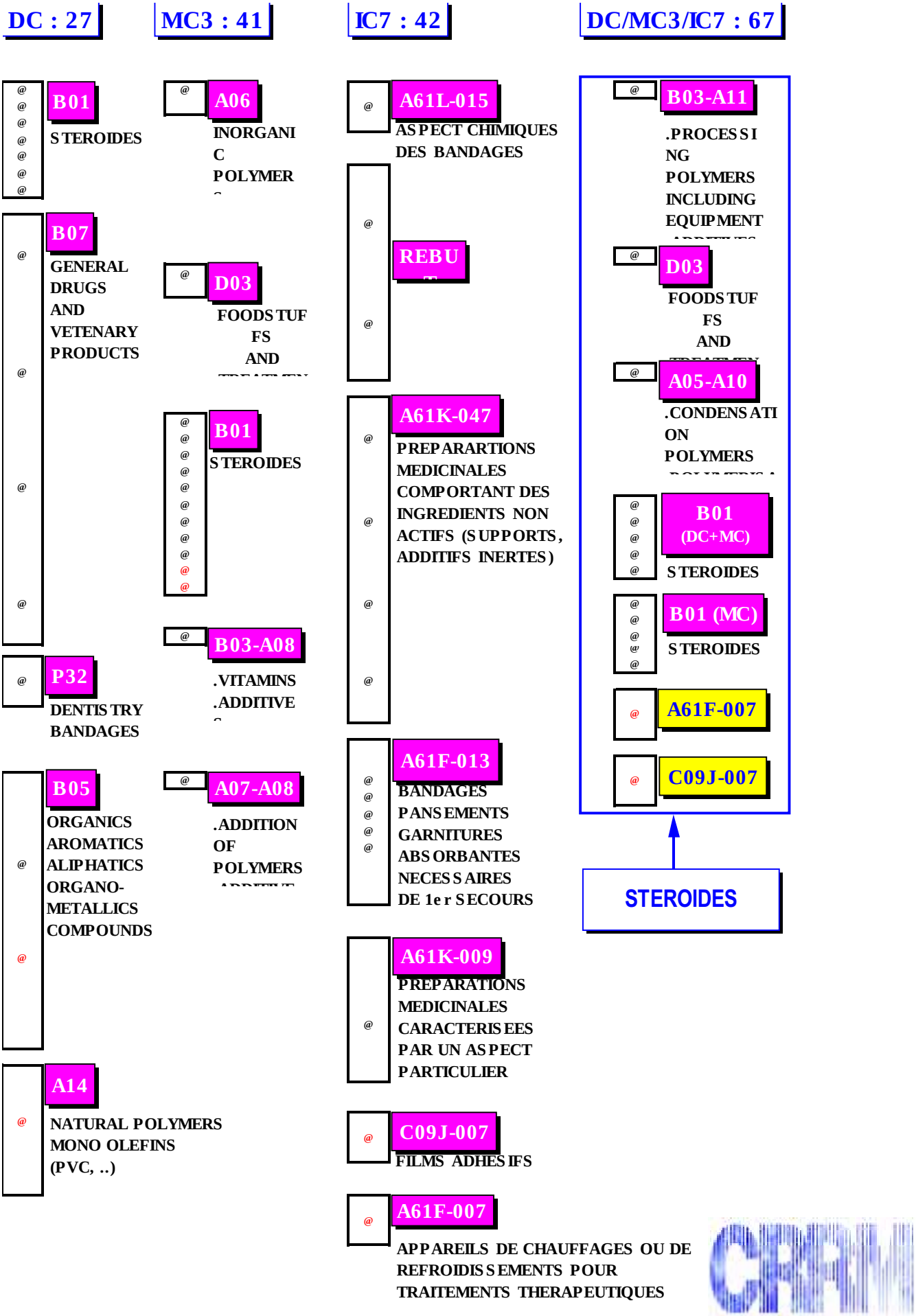
Pour montrer l'utilité de la combinaison des trois ensembles de codes, nous décrirons les résultats pour un exemple de sous-thème paraissant convaincant. Ce sous-thème correspond aux brevets revendiquant l'application de patchs transdermiques thérapeutiques pour la diffusion de composés actifs de la familles des stéroïdes, soit 15 brevets.

Comparons la répartition des brevets de ce sous-thème parmi les classes obtenues, pour les quatre matrices construites. La disposition des 15 brevets dans leurs classes d'appartenance est présentée symboliquement sur l'*illustration 7*.

Comment lire cette illustration?
Pour la matrice des DC à 3 caractères, les 15 brevets se distribuent dans 5 des 27 classes créées. Pour la matrice des MC à 3 caractères, ils se distribuent dans 5 classes parmi les 41... etc...
Sur cette illustration est indiqué, à côté de chaque classe, le code qui a le plus contribué à la construction de la classe. Ce sont les fameuses *étiquettes* décrites précédemment. Elles indiquent que le thème est spécifique aux brevets de cette classe et qu'il est pratiquement absent dans les brevets des autres classes.

Que nous dit cette illustration?
- Pour l'analyse faite à partir de la matrice des DC, une classe libellée *stéroïde* rassemble la moitié des brevets, les autres étant répartis dans d'autres spécialités. Donc, les codes DC ne décrivent correctement que la moitié des brevets *stéroïdes*.
- Pour l'analyse MC3, une classe *stéroïde* est constituée de 11 brevets; 2 autres brevets sont dans des classes singletons; et les 2 derniers appartiennent à de toutes petites classes. Ici, les MC3 ont très bien regroupé ces brevets dans des classes très homogènes.
- Pour l'analyse IC7, il n'existe aucune classe spécifique aux brevets *stéroïdes* et les brevets sont disséminés dans 7 classes de grandes tailles. Les codes IC7 ne permettent pas de regrouper les brevets concernés par les composés stéroïdes.

- Par contre, l'analyse réunissant les trois ensembles a parfaitement bien caractérisé ces brevets (une partie de cette matrice résultante est présentée par l'*illustration 8*). Ils sont tous dans des classes très homogènes. Les deux plus importantes de ces classes sont spécifiques aux codes descripteurs *stéroïde.* L'une est caractérisée par la co-présence des deux codes B01 qui représente les composés stéroïdes pour les deux codifications MC3 et DC, l'autre par la simple présence du B01 de la codification MC3. Cinq brevets *stéroïdes* se retrouvent isolés dans


DC : 27
MC3 : 41
IC7 : 42
DC/MC3/IC7 : 67
B01
STEROIDES
B07
GENERAL DRUGS AND VETENARY PRODUCTS
P32
DENTISTRY BANDAGES
B05
ORGANICS AROMATICS ALIPHATICS ORGANO-METALLICS COMPOUNDS
A14
NATURAL POLYMERS MONO OLEFINS (PVC, ..)
A06
INORGANIC POLYMER
D03
FOODSTUFFS AND
B01
STEROIDES
B03-A08
.VITAMINS .ADDITIVE
A07-A08
.ADDITION OF POLYMERS
A61L-015
ASPECT CHIMIQUES DES BANDAGES
REBU
A61K-047
PREPARARTIONS MEDICINALES COMPORTANT DES INGREDIENTS NON ACTIFS (SUPPORTS, ADDITIFS INERTES)
A61F-013
BANDAGES PANSEMENTS GARNITURES ABSORBANTES NECESSAIRES DE 1er SECOURS
A61K-009
PREPARATIONS MEDICINALES CARACTERISEES PAR UN ASPECT PARTICULIER
C09J-007
FILMS ADHESIFS
A61F-007
APPAREILS DE CHAUFFAGES OU DE REFROIDISSEMENTS POUR TRAITEMENTS THERAPEUTIQUES
B03-A11
.PROCESSING POLYMERS INCLUDING EQUIPMENT
D03
FOODSTUFFS AND
A05-A10
.CONDENSATION POLYMERS
B01 (DC+MC)
STEROIDES
B01 (MC)
STEROIDES
A61F-007
C09J-007
STEROIDES

CLASSES	REFERENCES	B12	dc_B07	B04	dc_A96	A12	dc_P34	A61K-009	B10	A61L-015	A61K-031	D09	dc_D22	dc_P32	A61K-047	A61F-013	B07	B11
4	17	X	.	X	.	.	.	.	X	.	X	.	.	.	.	.	X	.
4	25	X	X	.	.	.	.	.	X	.	.	.	.	.	X	.	.	.
4	42	X	.	.	.	.	.	.	.	.	X	.	.	.	.	.	.	.
4	44	X	.	X	.	.	.	.	X	.	X	.	.	.	.	.	X	.
4	75	X	.	.	.	.	X	.	X	.	.	.	.	.	.	.	.	.
4	117	X	.	.	X	X	.	X	X	.	X	.	.	.	.	.	X	.
4	145	X	.	X	.	.	.	.	X	.	X	.	.	.	.	.	.	.
6	6	X	X	X	X	X	.	X	X	.	.	.	.	.	.	.	X	X
6	33	X	X	X	X	X	X	X	.	X	.	X	X	.	X	.	.	.
6	57	X	X	X	X	X	.	.	X	.	X	X	X	X	.	X	.	.
6	76	X	X	X	.	.	.	.	X	.	.	X	X	X	.	X	X	.
6	140	X	.	X	.	.	X	X	X	X	X	.	.	X	X	X	.	X
7	3	X	X	X	X	X	.	.	.	.	.	X	X	X	.	X	.	X
7	4	X	X	.	.	.	.	.	.	.	.	X	X	X	.	X	.	.
7	43	X	X	X	X	X	.	.	X	.	.	.	.	X	.	X	.	.
7	67	X	X	X	.	.	.	.	X	.	X	.	.	.	X	.	.	.
7	146	X	.	X	.	.	X	.	X	X	X	.	.	.	.	.	.	X
#	38	X	X	X	X	X	X	X	.	X	X	.	.	X	.	X	.	.
#	45	X	X	X	X	X	X	.	.	X	.	X	X	.	.	.	.	.
#	58	X	.	X	X	X	X	X	X	X	X	.	.	.	X	.	.	.
#	118	X	X	.	.	.	X	X	.	X	.	.	.	X	.	X	.	X
#	138	X	X	X	.	.	X	X	X	X	.	X	X	X	X	.	X	.
#	26	X	X	X	X	X	X	X	.	X	.	.	.	.	X	.	.	.
#	13	X	X	X	X	X	X	.	.	X	.	X	X	.	.	.	.	.
#	142	X	.	X	X	X	X	.	X	X	X	.	.	.	.	.	.	.

CLASSES	REFERENCES	dc_B05	B06	A61M-037	dc_B03	B01	dc_B02	C12	A04	dc_S05	dc_D21	dc_C03	S05	A61N-001	A61M-035	C04	dc_B04	dc_A14
4	17	X	X	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	25	X	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	42	X	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	44	X	X	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	75	X	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	117	X	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	145	X	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	6	.	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
6	33	.	X	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
6	57	.	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
6	76	.	X	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
6	140	X	X	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
7	3	.	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
7	4	.	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
7	43	.	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
7	67	X	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
7	146	.	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
#	38	.	.	X	.	X	.	.	.	.	.	.	.	.	.	.	.	X
#	45	.	.	.	.	.	.	.	X	.	.	.	.	.	.	.	.	.
#	58	X	.	X	.	.	.	.	.	.	.	.	.	.	.	.	.	.
#	118	.	.	.	.	.	.	.	.	.	.	.	.	.	X	.	.	.
#	138	X	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
#	26	.	.	.	.	X	.	.	X	.	.	.	.	.	.	.	.	.
#	13	.	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.
#	142	.	.	.	.	X	.	.	X	.	.	.	.	.	.	.	.	.

CLASSES	REFERENCES	dc_P73	dc_P31	B02	dc_B01	D08	A06	A05	dc_A26	dc_P33	A61K-037	B05	A61B-005	dc_G03	G03	B03	A11	A10
4	17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	42	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	75	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	117	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	145	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	76	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	140	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
7	3	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.	.
7	4	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.	.
7	43	.	.	.	X	.	X	.	.	.	.	.	.	.	.	.	.	.
7	67	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.	.
7	146	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.	.
#	38	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
#	45	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
#	58	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
#	118	.	.	.	.	.	.	.	.	X	.	.	.	.	.	.	.	.
#	138	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
#	26	.	.	X	.	.	.	.	.	.	.	.	.	.	.	X	X	.
#	13	.	.	.	X	.	.	.	.	.	.	.	.	X	X	.	.	.
#	142	.	.	.	X	.	.	X	.	.	.	.	.	.	.	.	.	X

CLASSES	REFERENCES	A09	A24F-047	dc_P15	C09J-007	B32B-007	C11	C10	A08	A07	A61J-003	dc_S03	dc_A97	C08K-005	A61F-007	dc_P24	C08G-067	D03	Year	Name
4	17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	90	CAMP
4	25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	89	NIHA
4	42	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	89	NOVA
4	44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	88	BERC
4	75	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	88	ELBA
4	117	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	85	SCHE
4	145	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	79	MOME
6	6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	87	KIMY
6	33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	89	RIKL
6	57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	86	RUTF
6	76	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	84	YAMA
6	140	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	83	AMHP
7	3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	90	ROBE
7	4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	90	ROBE
7	43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	89	PHAR
7	67	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	88	EAST
7	146	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	79	KEYP
#	38	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.	.	89	LOHM
#	45	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.	.	89	BADI
#	58	.	.	.	X	.	.	.	.	.	.	.	.	.	.	.	.	.	88	LOHM
#	118	.	.	.	.	.	.	.	.	.	.	.	.	.	X	.	.	.	86	CIBA
#	138	.	.	.	.	.	.	.	.	.	.	.	.	.	X	.	.	.	84	YAMA
#	26	.	.	.	.	.	.	.	X	.	.	.	.	.	.	.	.	.	89	NITL
#	13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	X	90	NOVE
#	142	.	.	.	.	.	.	.	X	X	.	.	.	.	.	.	.	.	79	KEYP

des classes ou presque seuls. Cet isolement est du au fait qu'ils contenaient d'autres codes très rares. Ces codes rares les ont marginalisés du reste du sous-thème. Deux de ces brevets ont été placés dans deux classes dont la spécificité est créée par un code IC7. Ces deux codes sont rares dans les brevets car ils ont des revendications très accentuées sur un caractère atypique: le code C09J-007 revendique une grande qualité d'adhésion, le code A61-007 concerne des patchs dont le composé actif se libère de la matrice de diffusion par différence de température avec la peau.

Donc grâce à l'utilisation des trois ensembles de codes réunis, l'agrégation établie fait ressortir simultanément les caractéristiques décrites par les différents ensembles de codes:

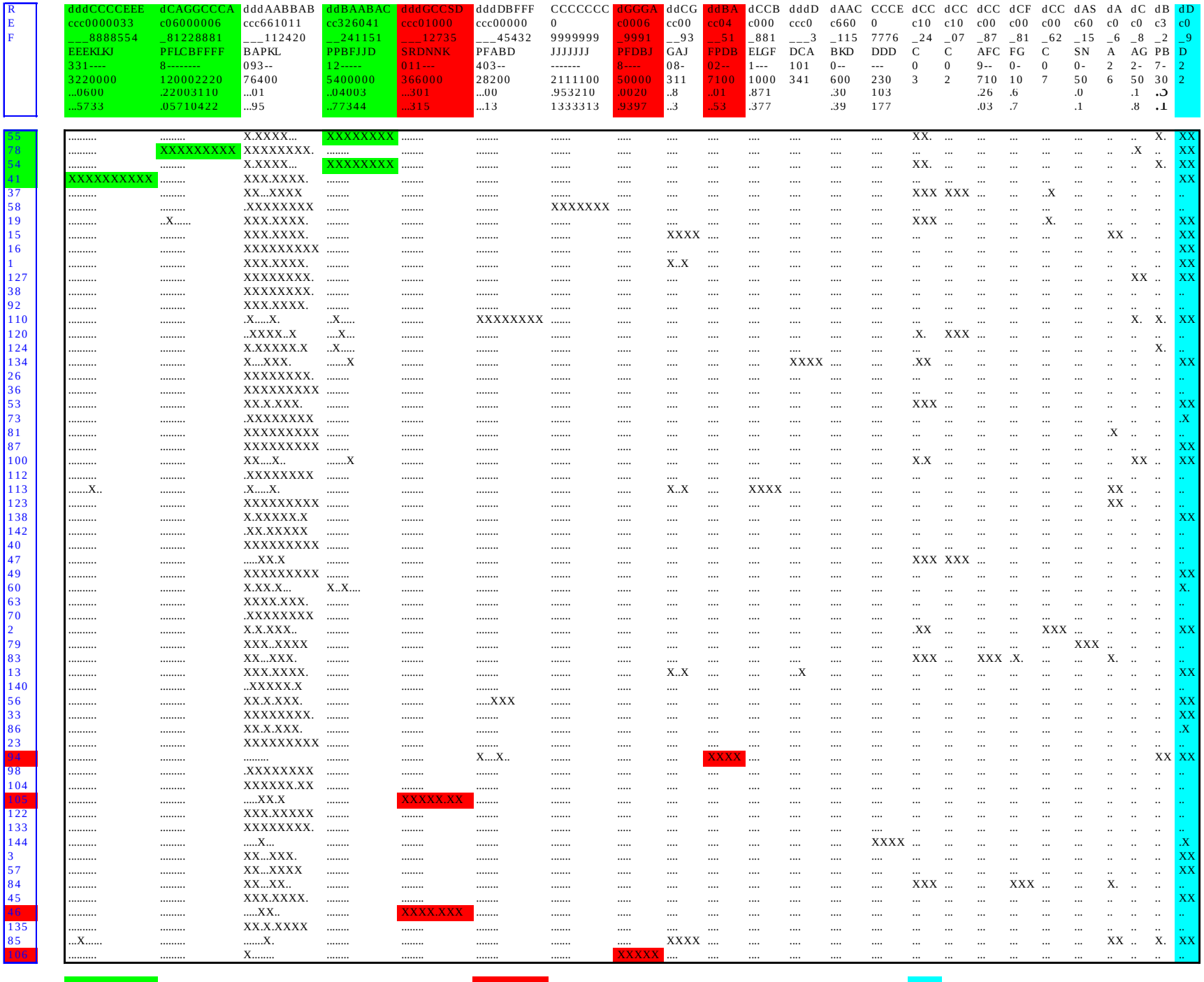
- le caractère *composés stéroïdes* décrit par les codes DC et MC3 et qui disparaît dans l'analyse de la matrice des codes IC7
- les caractères spécifiques plus rares uniquement décrits par la codification plus fine IC7 et qui sont noyés dans d'autres thèmes dans les analyses des matrices des codes DC et MC3

Etude des correspondances entre codes

Les résultats obtenus par cette analyse s'ils n'ont pas répondu à toutes nos espérances, ont toutefois révélé des informations inattendues et très prometteuses.

Nous espérions pouvoir faire apparaître les synonymies et les déficiences entre les trois ensembles de codes. L'analyse n'a pas livré ces correspondances entre les ensembles de codes, probablement parce que le nombre de brevets de notre échantillon était trop restreint. Au final, seuls les codes, soit très rares, soit présents presque partout, ont été regroupés dans des classes de tailles raisonnables. Les codes entre ces deux extrêmes, c'est-à-dire les codes dont les fréquences ne sont ni fortes ni faibles, ont créé une multitude de petites classes sans cohésion entre elles. Or, ce sont précisément ces codes qui sont porteurs de l'information la plus intéressante puisque les deux autres corps de codes se rapportent pour le premier au bruit de fond et pour le second à l'information triviale. Pour une taille de corpus bien plus élevée, la plage intermédiaire de codes porteurs d'information augmenterait vers des valeurs de fréquences plus élevées et, les poids des codes grandissant, des agrégats pourraient se former.

Cette analyse nous a tout de même apporté des satisfactions. Le fait que les codes très rares et toujours présents ensemble soient très bien discriminés n'est pas uniquement un inconvénient.Comme on le voit sur l'*illustration 9,* il est très facile de détecter les brevets


R
E
F
dddCCCCEEE ccc0000033 ___8888554 EEEKLKJ 331---- 3220000 ...0600 ...5733
dCAGGCCCA c06000006 _81228881 PFLCBFFFF 8-------- 120002220 .22003110 .05710422
dddAABBAB ccc661011 ___112420 BAPKL 093-- 76400 ...01 ...95
ddBAABAC cc326041 __241151 PPBFJJD 12----- 5400000 ..04003 ..77344
dddGCCSD ccc01000 ___12735 SRDNNK 011--- 366000 ...301 ...315
dddDBFFF ccc00000 ___45432 PFABD 403-- 28200 ...00 ...13
CCCCCCC 0 9999999 JJJJJJJ ------- 2111100 .953210 1333313
dGGGA c0006 _9991 PFDBJ 8---- 50000 .0020 .9397
ddCG cc00 __93 GAJ 08- 311 ..8 ..3
ddBA cc04 __51 FPDB 02-- 7100 ..01 ..53
dCCB c000 _881 ELGF 1--- 1000 .871 .377
dddD ccc0 ___3 DCA 101 341
dAAC c660 _115 BKD 0-- 600 .30 .39
CCCE 0 7776 DDD --- 230 103 177
dCC c10 _24 C 0 3
dCC c10 _07 C 0 2
dCC c00 _87 AFC 9-- 710 .26 .03
dCF c00 _81 FG 0- 10 .6 .7
dCC c00 _62 C 0 7
dAS c60 _15 SN A 0- 50 .0 .1
dA c0 _6 A 2 6
dC c0 _8 AG 2- 50 .1 .8
dB c3 _2 PB 7- 30 .3 .1
dD c0 _9 D 2 2
55 78 54 41 37 58 19 15 16 1 127 38 92 110 120 124 134 26 36 53 73 81 87 100 112 113 123 138 142 40 47 49 60 63 70 2 79 83 13 140 56 33 86 23 94 98 104 105 122 133 144 3 57 84 45 46 135 85 106
DOCUMENTS ORIGINAUX
DOCUMENTS NON PERTINENTS
CORRESPONDANCE DE CODES

décrits par une famille de codes très rares. Ces brevets représentent deux catégories:

- des brevets non pertinents:
  ils ont très peu de codes dans la classe des codes "communs" à tous les brevets (la 3ème classe)

- des brevets pertinents dont les revendications sont très originales:
  ils possèdent pratiquement tous les codes de la classe de codes "communs".
  Quelques exemples de ces brevets originaux:
  - La référence 41: utilisation d'oxydes métalliques dans un précipité aqueux qui peut être moulé et mis en forme pour rentrer dans la constitution de patchs.
  - La référence 78: polymère ayant des caractéristiques de très grande transparence, haute perméabilité, flexibilité et doux en état hydraté. Outre les applications en comme membrane de séparation de gaz, il peut servir de garniture de blessure perméable à l'oxygène, de lentille de contact, de patch buccal ou d'implant dans le corps.
  - La référence 55: patch qui délivre un haute dose initiale de médicament suivi d'une dose basse et régulière (système à plusieurs réservoirs d'agent actif).

● CONCLUSION

Cette méthodologie nous a permis de mettre en évidence plusieurs caractéristiques:

Grâce à la complémentarité des codes :

On relève des détails techniques très précis qui sont en général noyés au travers d'aspects très généraux et par conséquent très difficiles à détecter; mais ceci sans perdre de vue les aspects généraux auxquels ils se rapportent.

Pour l'exemple du sous-thème *principes actifs de type stéroïde*

* l'utilisation d'un seul type de codification permet:
  - soit de connaître plus ou moins rapidement quels sont les documents qui traitent des stéroïdes de façon générale, sans aller beaucoup plus loin dans le détail.

  - soit de déterminer plus ou moins rapidement leurs aspects spécifiques, sans savoir qu'ils parlent de stéroïdes.

* L'utilisation simultanée de différents types de codifications permet non seulement de déterminer très facilement les aspects généraux traités dans les documents (ici l'aspect stéroïde), mais aussi de révéler les aspects spécifiques développés dans ces mêmes documents (par exemple des problèmes de température ou d'adhésion).

Donc, choisir d'exploiter plusieurs champs de descriptions pour un traitement bibliométrique permet d'aboutir à une **MEILLEURE CARACTERISATION** du corpus

Grâce à la correspondance des codes :

On met en regard des ensembles de codes qui sont fortement dépendants (ou proches) les uns des autres parce que très souvent employés conjointement dans les références. Ceci nous a permis de:

- déceler très aisément les documents non pertinents de notre corpus c'est à dire ceux qui représentent le **BRUIT**
- reconnaître très facilement les **DOCUMENTS ORIGINAUX** (au sens innovateur du terme) qui sont basés sur l'emploi ou la description de techniques, de méthodes, de procédés qui se démarquent des autres documents.

La représentation de l'information obtenue par les outils d'analyses

L'expertise et l'interprétation des résultats ne sont fonction que de l'information qui résulte des traitements analytiques auxquels on a recours.

De ce fait, afin que l'ensemble des experts puisse interpréter de façon objective les résultats, il est nécessaire d'utiliser des méthodes de traitements qui permettent non pas de résumer l'information de départ, mais de la restructurer afin d'en dégager les faits marquants sans perdre le reste. Trop souvent négligé par les méthodes d'analyses traditionnelles, cet ensemble rebut constitue une part considérable de l'information initiale (lois de Zipf- Bradford). Bien que ces éléments soient présents à des fréquences très faibles, ils n'en contiennent pas moins une information intéressante. En effet, on y relève **L'INFORMATION MARGINALE** qui selon le cas peut se traduire en termes d'information **INNOVANTE** ou **DISCORDANTE** (le bruit).

L'ANALYSE RELATIONNELLE parait de ce fait parfaitement appropriée à l'analyse bibliométrique. Elle **NE NEGLIGE AUCUNE INFORMATION** même si sa faible présence lui donne a priori un caractère mineur.

Le rapport temps d'analyse - temps d'expertise

Afin que les experts du domaine puissent se consacrer pleinement à l'expertise du sujet au travers des résultats livrés par le traitement, celui-ci doit prendre le moins de temps possible.
Pour minimiser ce rapport temps d'analyse - temps d'expertise, on se doit d'utiliser des méthodes de traitement, de calcul, d'analyse, qui permettent d'obtenir des résultats dans des délais extrêmement brefs et qui permettent de s'utiliser de façon systématique afin de réorienter les analyses selon les interprétations que l'on obtient.
C'est la première caractéristique qu'un système de surveillance doit vérifier pour assurer un fonctionnement performant. L'accélération de l'obsolescence des technologies impose à la veille technologique d'être le pourvoyeur d'une **INFORMATION ELABOREE DANS DES TEMPS RESTREINTS**. Pourquoi dans ces conditions ne pas profiter des avantages que nous offre l'informatique pour le traitement des données. C'est dans cet état d'esprit que le logiciel **DATAVIEW** a été conçu.

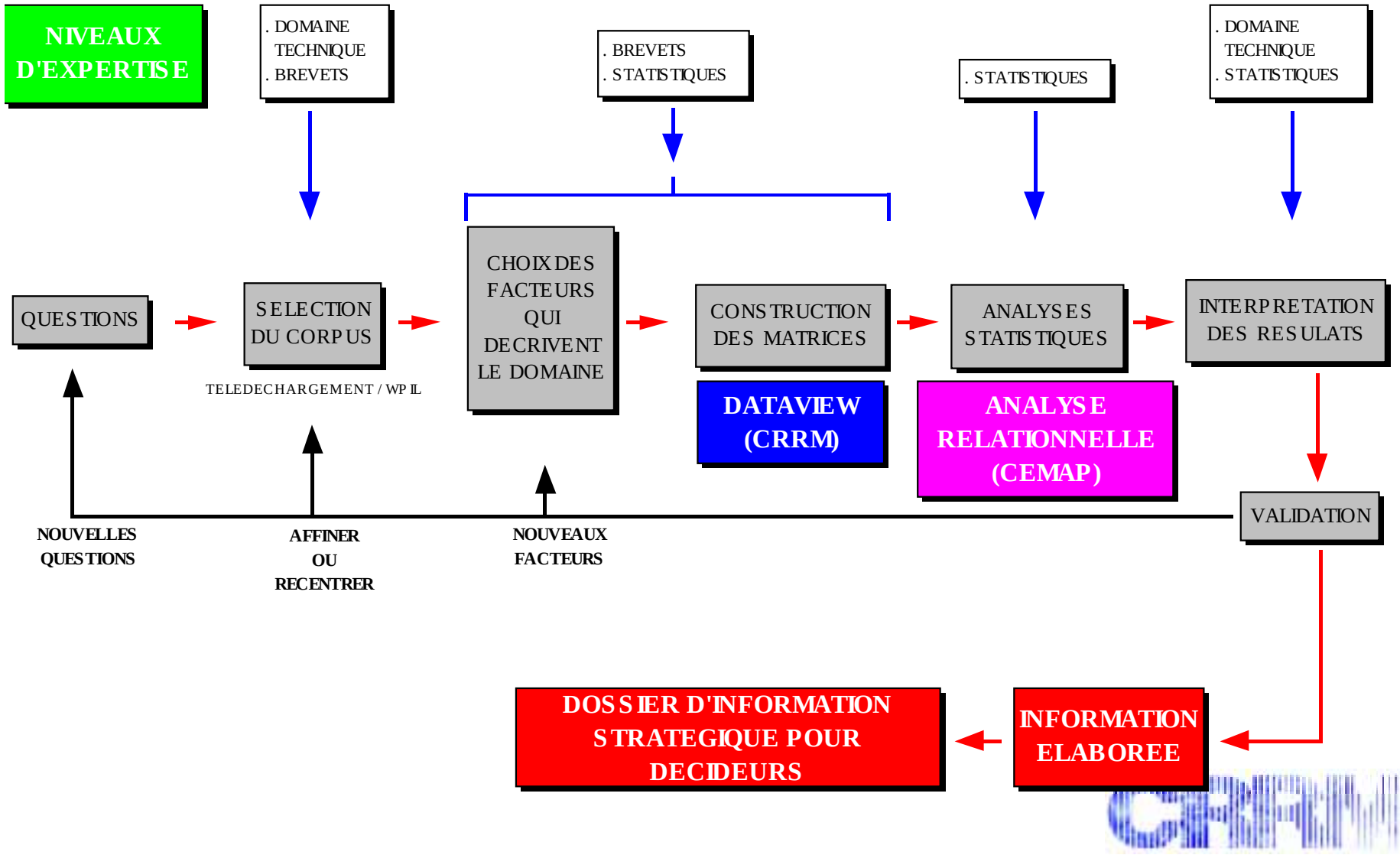
Le recours permanent à des experts différents

Pour que l'analyse délivre une **INFORMATION FIABLE** à but stratégique, il est indispensable de valider chaque étape dans l'élaboration du dossier, depuis la sélection du corpus jusqu'à l'interprétation des résultats, par des **NIVEAUX D'EXPERTISES** différents et adaptés.

- Expert du domaine technique étudié
- Expert brevet
- Expert de l'information
- Expert en statistiques

Ce dernier point est illustré sur le synoptique des traitements effectués pour cette étude (*illustration 10).*
Ce synoptique est volontairement replacé dans le contexte d'un veille technologique industrielle. Le processus démarre sous la manifestation de questions sur des sujets sensibles à l'entreprise (facteurs critiques). L'aboutissement de la chaîne du traitement est l'élaboration de dossiers stratégiques pour informer les décideurs de la situation présente et des évolutions de tendances.
Ce synoptique montre les exigences et les compétences qu'impose l'insertion de l'outil bibliométrique dans un système de surveillance industriel. Ce sont les contraintes que doit respecter un système de veille pour permettre de traiter des sujets dont la masse et la complexité des connaissances ne peuvent être appréhendés par de simples traitements manuels, dans des temps acceptables.

SYNOPTIQUE


NIVEAUX D'EXPERTISE
. DOMAINE TECHNIQUE
. BREVETS
. BREVETS
. STATISTIQUES
. STATISTIQUES
. DOMAINE TECHNIQUE
. STATISTIQUES
QUESTIONS
SELECTION DU CORPUS
TELEDECHARGEMENT / WPIL
CHOIX DES FACTEURS QUI DECRIVENT LE DOMAINE
CONSTRUCTION DES MATRICES
DATAVIEW (CRRM)
ANALYSES STATISTIQUES
ANALYSE RELATIONNELLE (CEMAP)
INTERPRETATION DES RESULATS
VALIDATION
NOUVELLES QUESTIONS
AFFINER OU RECENTRER
NOUVEAUX FACTEURS
DOSSIER D'INFORMATION STRATEGIQUE POUR DECIDEURS
INFORMATION ELABOREE

BIBLIOGRAPHIE

[1] La veille technologique
sous la direction de H Dou, H Desval
Dunod, 1992

[2] Pratique de la veille technologique
F Jakobiak
Les éditions d'organisation, 1991

[3] L'analyse des données au service de la bibliométrie. Outils de veille technologique à la dimension des moyennes entreprises.
H Dou, L Quoniam, H Rostaing, W Nivol
Revue Française de biliométrie, Vol 8, p 27-67, déc. 1990

[4] The relationship of information science et the social sciences- a cocitation analysis
H Small
Information porcessing & management, Vol 17, N• 1, p39-50, 1981

[5] L'analyse des associations
B Michelet
Thèse de doctorat, Université de Paris VII, 26 oct 1988

[6] An analysis of citations in statistical journals
K Fz Agirre, J M Piris, F Tusell
Proceedings of the first international symposium on applied stochastic models and data analysis, p 15-24, 23-26 april 1991, Granada, Spain

[7] Application de l'analyse relationnelle à la veille technologique: des outils d'analyse de l'information documentaire
C Bédécarrax, C Huot
Revue française de bibliométrie, Vol 9, p 66-80, sept 91