

**COMMENT MEDIATISER LES ANALYSES
BIBLIOMETRIQUES.
PRESENTATION DE LA RECHERCHE EN CHIMIE A
MARSEILLE SUR 10 ANNEES.**

Henri Dou, Luc Quoniam, Parina Hassanaly, Jeanne Pullino

CENTRE DE RECHERCHE RETROSPECTIVE DE MARSEILLE
C.R.R.M. Centre Scientifique de St Jérôme
13397 Marseille cédex 13
91029094 vidéotex 36.16 code CRRM

Résumé: la production d'indicateurs scientifiques pose plusieurs problèmes: sélection des corpus et des bases de données d'où les productions de recherche seront extraites et traitées, méthodes d'analyse, formulation des résultats, présentations de ces derniers aux chercheurs ou aux décideurs.

Le langage, la formulation, le traitement avec personnification ou non, posent de multiples problèmes dans un pays, où la publication scientifique reste encore un des seuls critères de jugement d'un chercheur.

Nous présenterons, en prenant comme thème la chimie à Marseille sur dix années, à partir de l'analyse des données fournies par les Chemical Abstracts, ces résultats en mettant en relief les formulations et les méthodes susceptibles de faire accepter sans polémiques ces études.

1 - Situation du problème:

Dans un article récent publié dans La Recherche (1), il a été mis en évidence que la corrélation directe entre argent investi et volume de publications n'étaient plus en corrélation dans certains Pays développés comme les USA, la Grande Bretagne, la RFA. Il est donc clair, que dans certains domaines sensibles, la publication ne représente ou ne représentera plus le vecteur de connaissances, tel qu'il est encore perçu par la communauté scientifique française.

A l'inverse, la France constitue un point singulier, où les carrières sont encore liées, dans l'Enseignement Supérieur et dans les Organismes de Recherche, à la publication scientifique, ce critère étant dans presque tous les cas exclusif.

Ce n'est donc pas par hasard, que l'on voit se développer en France une politique d'indicateurs, qui se voudraient sélectifs.

Ce développement ne va donc pas aller sans problèmes, puisque par la nature même de certains indicateurs (par exemple les Sciences Citations Index), la science est divisée en suiveurs et suivis, le nombre de ces derniers étant beaucoup plus faible que le précédent, mais pouvant varier selon les disciplines.

Il est donc clair, que plusieurs attitudes seront possibles, vis à vis d'une part de l'application des indicateurs, et d'autre part sur le choix même de ces derniers.

2 - Quels indicateurs:

Les personnes qui se trouvent placées devant la création d'indicateurs cohérents, peuvent avoir deux attitudes, qui sont étroitement liées à la méthodologie utilisée, et aux buts fixés. En effet, on peut considérer une production scientifique de deux manières:

* dans son ensemble, en admettant que tout travail publié mérite considération, puisqu'il conduit à l'exposé de résultats, avalisés par des pairs (les arbitres). Cette attitude est d'autant plus justifiée, que si on se place au plan de la naissance des innovations, il n'est pas dit que celles-ci aient un lieu privilégié pour éclore. Dans le cas de la surveillance technologique, la situation est plus grave encore, puisque tout travail publié est opposable à une demande de prise de licence.

* dans sa partie dite la "plus noble", c'est à dire dans le groupe des revues fondamentales les plus citées. Ceci conduit à émettre a priori un jugement sur la qualité d'un travail qui pour être satisfaisant doit être publié dans une liste de revues type. En outre cela revient à écarter les brevets, qui dans certains domaines véhiculent plus de 60% d'informations originales (2). Cette analyse, repose, en exclusivité sur la base de données des citations index, développée par E. Garfield (3).

3 - La différence entre les deux méthodes:

Il est donc évident que les deux méthodes son différentes et dans certains cas complémentaires, car il serait naïf de croire que la science ne peut avancer sans un travail de fond, parfois obscur, ou une succession d'applications techniques. Les citations, leurs volumes, leurs agrégats, étant des données reliées à l'avancement des sciences, par "créations de concepts successifs".

Ceci induit des choix très différents au niveau des méthodes de travail. En effet, l'étude des productions globales, va nécessiter l'utilisation de bases de données les plus larges possibles dans un domaine défini avec si possible la prise en compte des brevets. La méthode des citations ne reposera que sur une base de données unique, celle des citations index.

4 - La sensibilité des chercheurs:

Cette sensibilité, est exacerbée en France, par le jeu des commissions et des critères de "jugement". Il est évident, que mettre en regard de la production des chercheurs, un indice de citation dont on sait qu'il est relativement faible en moyenne quelle que soit la situation, est dangereux.

D'autre part, une connaissance des indicateurs du type "citation" met en évidence que la division entre production de base (la plus citée) et production courante (qui est celle qui cite la précédente), revient à diviser la production de recherche en 20% de relativement bon et 80% qui peuvent être considérés comme relativement moyens, voire à négliger. On comprend aisément l'utilisation déviée qui peut être faite de tels résultats.

D'autre part, il est aussi évident que les problèmes de langue, de paiement des publications en devises fortes, de l'exclusion des brevets, de la rediffusion des résultats, va jouer à plein dans un tel système.

Pour éviter ces écueils, et les polémiques nous avons préféré, en un premier temps, nous concentrer sur des méthodes applicables à des analyses globales.

En effet, il est souhaitable de situer la situation de la recherche dans un domaine précis et dans un pays, puis de réaliser des comparaisons permettant de localiser les forces et les faiblesses. Cet état des lieux, ainsi que celui des vecteurs nationaux de diffusion, nous paraît fondamental, avant une approche plus fouillée. (*)

(*) Nous prenons de ce fait en compte l'aspect capitalisation des savoirs. Une publication représente le savoir scientifique d'un ou de plusieurs auteurs. Ne pas en tenir compte revient à ne pas capitaliser cet acquis.

5 - Quelles sources utiliser pour de telles analyses:

L'idéal serait de posséder une source d'information nationale, récapitulant en priorité les travaux nationaux, avec la possibilité d'extraire ces derniers par disciplines, villes, auteurs, ...

Nous ne possédons pas un tel outil. La base de données Pascal, pourrait le devenir, si une volonté nationale se faisait jour dans ce domaine.

Il est donc nécessaire soit de faire appel à une base de référence, qui sera, reconnue par l'ensemble des acteurs, ceci est le cas des Chemical Abstracts. Mais, ce choix, induit certaines limites, comme celle de la couverture de la base, de la distorsion des adresses, puisque c'est celle du premier auteur qui est pris en compte. Dans le cas de notre étude cela introduit une distorsion de l'ordre de 10%. (Distorsion mesurée à partir de la production globale, en Chimie d'une Unité Associée au CNRS).

Néanmoins, cette approche reste cohérente, car au niveau de la chimie par exemple, 14.000 sources différentes sont analysées par les Chemical Abstracts, alors que seulement 4000 sources environ sont prises en compte par les Citations Index, pour tous les domaines scientifiques. La comparaison de ces deux chiffres met bien en évidence une différence dans les méthodes, et dans les buts recherchés.

Le choix de la méthode même, qui va donner naissance à des indicateurs, est donc politique.

6 - La méthode utilisée:

Dans le cas de l'analyse qui sera présentée, la méthodologie utilisée a été décrite dans divers travaux (4,5,6). Elle fait appel au traitement du champ des sections des Chemical Abstracts, et à la constructions à partir de ces derniers de représentations des domaines de recherche et des réseaux de relations entre ces derniers.

7 - La médiatisation:

Nous appelons médiatisation, la présentation des résultats de ces études, qui mettent en jeu rappelés le un processus compliqué de sélection des corpus, télédéchargement, analyse automatique, vérification par des experts, puis mise en forme définitive.

*** que présenter:**

A notre avis, en l'état actuel des personnels de recherche en France, il semble nécessaire de fournir à ces derniers, aussi bien décideurs que chercheurs, une image concrète des productions, par zones géographiques. Nous avons choisi ici quelques villes, afin de mettre en valeur des différences.

Cette production doit aussi être située en volume, avec un calcul de la croissance ou de la décroissance, par zones, au niveau local et national. Rappelons que la même démarche (détermination des productions par villes a été faite très récemment par E. Garfield pour les USA). (7)

Ensuite, le suivi de ces productions est affiné, par son suivi dans le temps, ainsi que par la juxtaposition du réseau de relations entre les différents domaines analysés.

On constate que dans cette gradation des analyses, à aucun moment des noms de laboratoires, ni de chercheurs n'ont été cités. Seulement des domaines ont été situés, avec leur évolution dans le temps.

*** comment le présenter:**

Les ressources actuelles de l'informatique et la profusion des micro-ordinateurs, permettent de présenter de manière dynamique les résultats. A notre avis, il faut que la présentation soit accompagnée du minimum de renseignements permettant de comprendre clairement la méthode.

D'autre part, une attention particulière doit être prise, lorsque des mots décrivant des domaines sont utilisés. En effet, les mots n'ont pas la même signification pour tous, et l'omission de certains entraîne généralement la non validation ou le rejet du modèle par des communautés scientifiques dans leur ensemble. C'est pour cela que nous avons, en un premier temps fait appel à des codes, divisant le domaine de façon constante dans le temps en un certain nombre de parties. Celles-ci n'ont pas un rapport direct avec l'évolution fine des spécialités, mais représentent un ensemble logique choisi par le producteur de la base.

Ceci évite toute polémique quand au choix des mots, ou à leur apparition soudaine par regroupements ou choix des experts. Il va sans dire, que pour une analyse fine, il faudra aller au niveau des concepts, mais on travaillera alors sur des ensembles plus réduits, sélectionnés à partir des représentations précédentes.

8 - Les résultats :

Ils sont présentés sous forme de scénario, les logiciels spécialisés que nous avons développés et utilisés, donnent naissance à des "images" qui sont capturées puis programmées dans le temps suivant un ordre logique à partir de logiciels commerciaux de la classe de IBM ST. Le tout est fourni sous forme de disquettes lisibles par un seul appel au prompt, (écran EGA, nécessaire, couleur si possible).

Par exemple un jeu de deux disquettes permet de représenter l'évolution de la chimie à Marseille sur 10 ans, en donnant une image statistique cohérente par la seule vue des changements qui vont bien au delà en volume de la limite des erreurs possibles.

Les tests de cette présentation, effectués avec différents chercheurs, ont bien mis en évidence la méconnaissance par ces derniers de l'ensemble des recherches dans lequel s'insère leur travail, ainsi que celle du volume total de production directement relié à leur domaine.

Les quelques figures présentées donnent une simple idée des résultats, mais imparfaite du fait de l'absence de dynamique et de couleurs. (figures 1,2,3,4)

9 - Les discussions :

A notre avis, les données présentées permettent sans discussion de voir la direction des évolutions. D'autre part le réseau des recherches, permet de mettre en évidence si cela est nécessaire les laboratoires qui travaillent sur un axe précis (une arête du graphe, ou un noeud particulier).

A partir de cela on peut aller plus au fond, jusqu'au laboratoire aux auteurs, à leur localisation géographique. Dans le cas où on voudrait voir dans quelle mesure des liens nouveaux pourraient être créés à partir des noeuds existants, on peut entrer ces derniers dans la base de données des Chemical Abstracts et lister les travaux effectués (s'il y en a) puis en déduire des orientations possibles, etc... On peut aussi dans le cadre d'une base de données telle que les Chemical Abstracts, sélectionner un lien existant et voir qu'elle est l'activité scientifique qui en découle, au niveau des publications, des brevets.

10 - Conclusion :

Les analyses globales, présentées sous forme de scénario et disponibles au sein d'une communauté scientifique, devraient permettre aux chercheurs de mieux se situer dans l'ensemble des recherches effectuées au niveau de leur discipline, d'une zone géographique, etc...

Les commentaires doivent à notre avis se limiter à l'exposé de la méthodologie, de manière à ce qu'il n'y ait pas d'équivoque. Ensuite, nous pensons que les décideurs, les chercheurs, pourront dans la mesure de leur curiosité ou de leurs intérêts, utiliser les résultats.

C'est pour cela, qu'il ne nous paraît pas souhaitable au niveau de suggérer des interprétations. En effet, l'utilisation de ces évolutions doit nécessairement s'accompagner de la connaissance de celles des équivalents temps plein chercheurs, des potentiels de recherche engagés, des crédits disponibles On entre alors dans le cadre de la programmation de la

recherche, qui n'est pas l'objet de l'exposé méthodologique présenté dans ce travail.

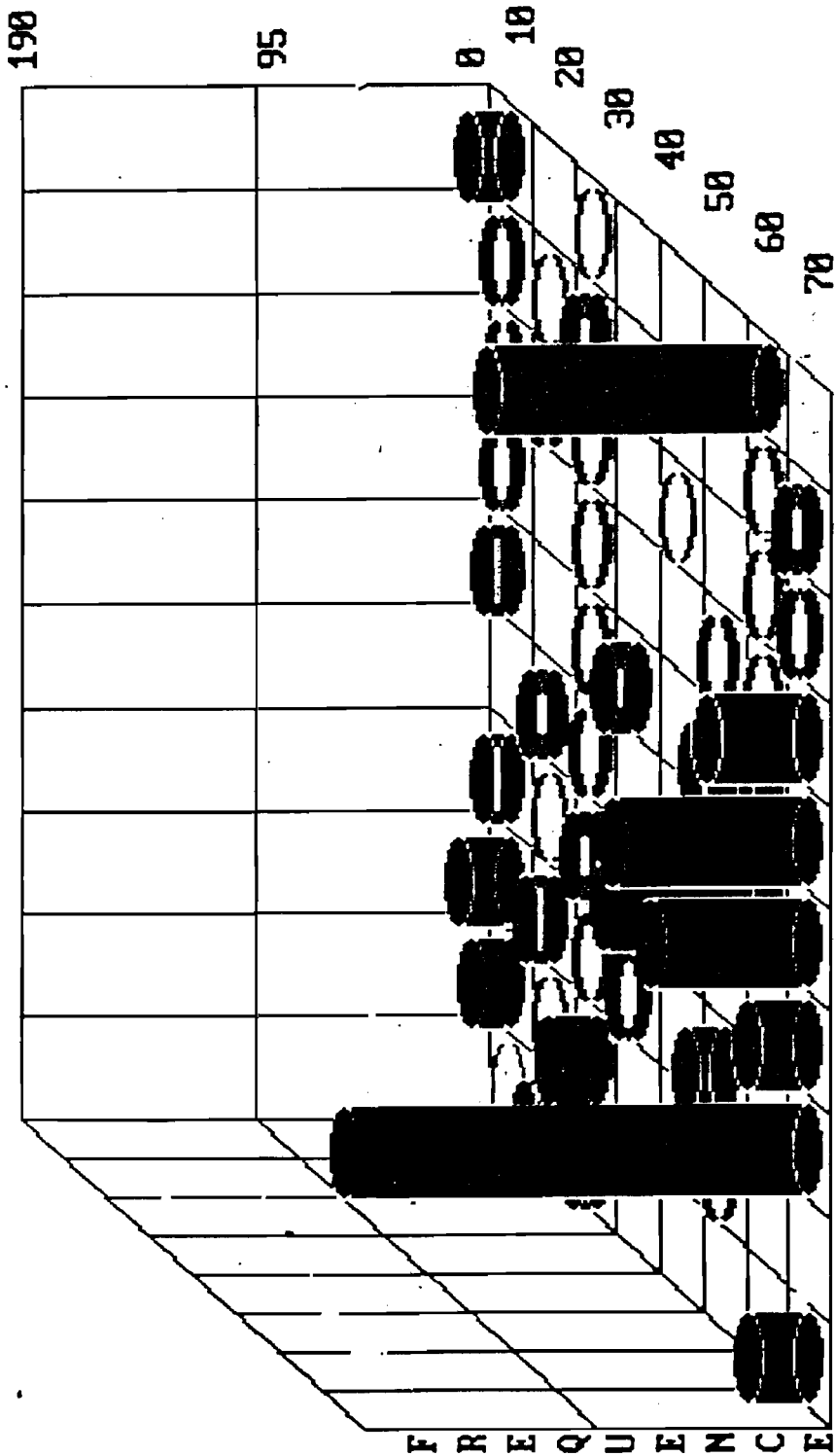
Nous préférons, en un premier temps essayer de donner une vision saine de la production de recherche, afin de permettre le développement de discussions constructives, et non de donner naissance à une polémique basée sur la "qualité" des travaux produits. Toute publication, acceptée dans une revue scientifique découle d'un travail, plus ou moins brillant, donc d'un budget, ou d'une volonté de production. Elle fait donc partie d'un tout, et ne doit pas être négligée.

Bibliographie :

- 1 - La Science une Nouvelle Marchandise
FEME G.
La Recherche, v.20, p.429-435, 1989
- 2 - A unique source of information.
TERAPANE J. F.
Chemtech, v.8, p.272-278, 1978
- 3 - Mapping the World of biomedical engineering
GARFIELD, G.
Ann. Biomed. Engng. v.14, 97-108, 1986
- 4 - Mapping the Scientific Network of Patent and Non-patent Document for a Fast Scientific Analysis.
DOU H., HASSANALY P.
World Patent Information v. 10, p.133-149, 1988
- 5 - Teaching Bibliometric Analysis and MS/DOS Commands
DOU, H., QUONIAM, L., HASSANALY, P.
Education for Information v. 6, p. 411-423, 1988
- 6 - Information stratégique en recherche fondamentale.
Typologie et réseaux de relations entre domaines.
DOU, H., HASSANALY, P., QUONIAM, L., BLANC, R.,
PULLINO, J.
Revue Française de Bibliométrie, p.49-85, 1988
- 7 - Top 20 Science Cities in the US
GARFIELD E.
The Scientist, October 17, p.15, 1988

Pole principal 73

73 optical; electron; and mass spectroscopy and other related properties



LOCALISATION DES SECTIONS

91-029094 ou 36.16 code CRRM

ANALYSES BIBLIOMETRIQUES,

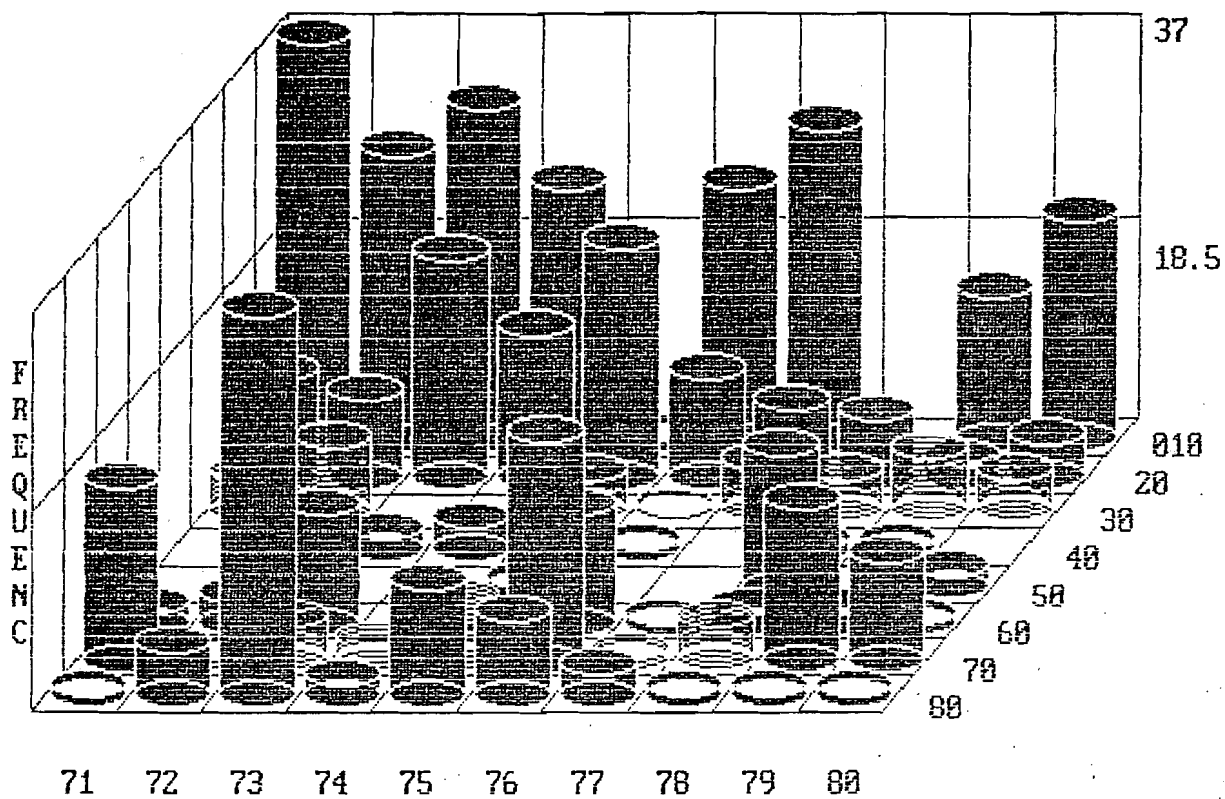
CENTRE DE RECHERCHE RETROSPECTIVE DE MARSEILLE,

Centre de St Jérôme, Université Aix - Marseille
13397 Marseille cédex 13

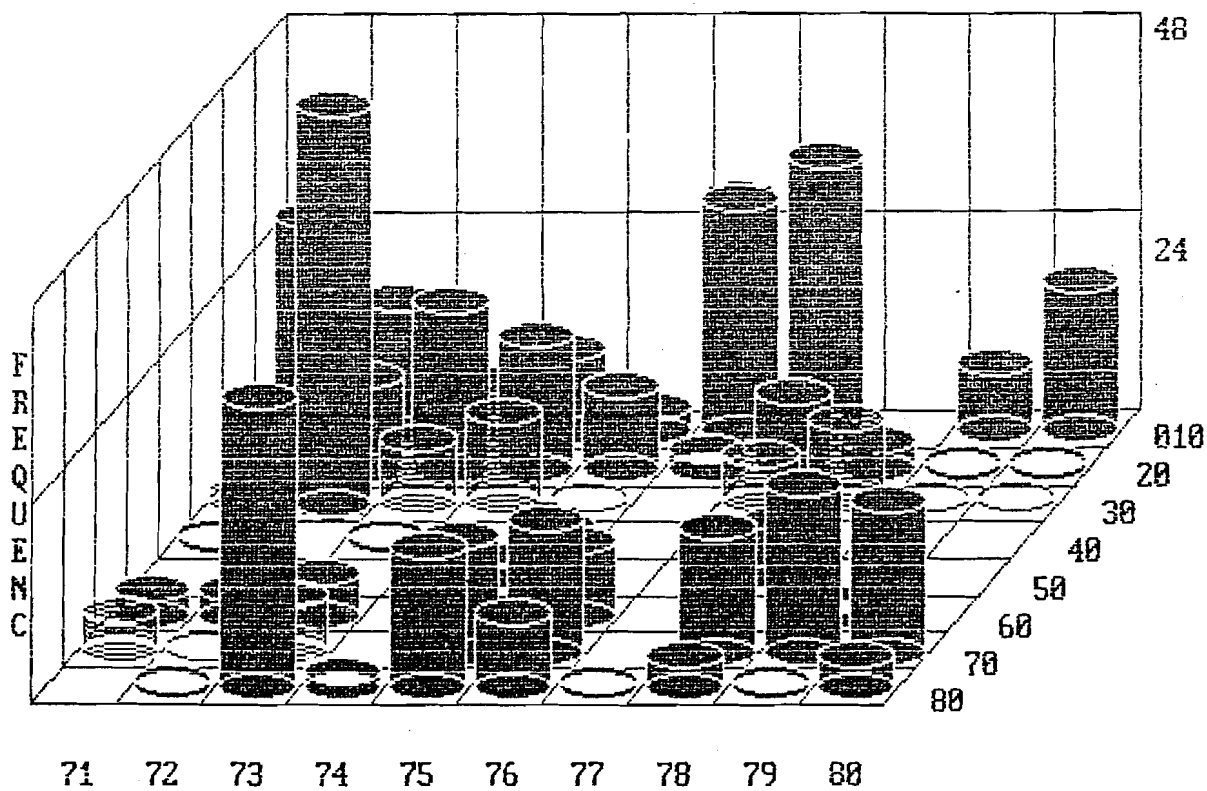
Orsay 1985 - seuil 2

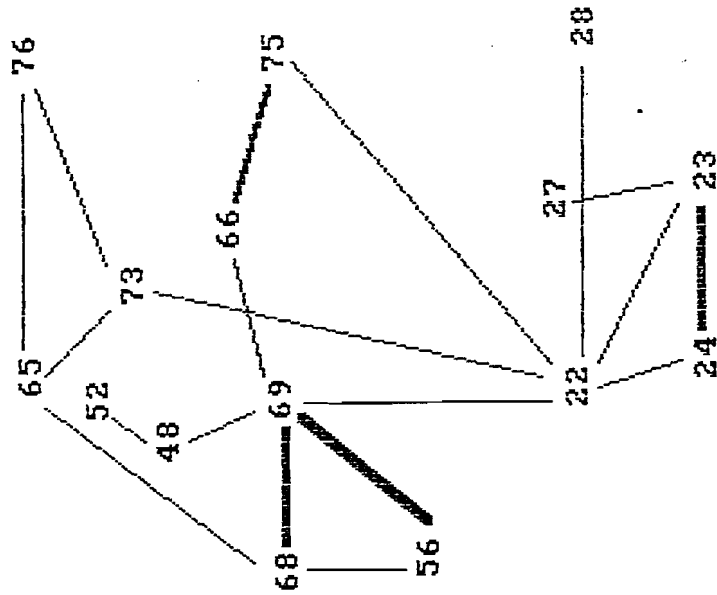
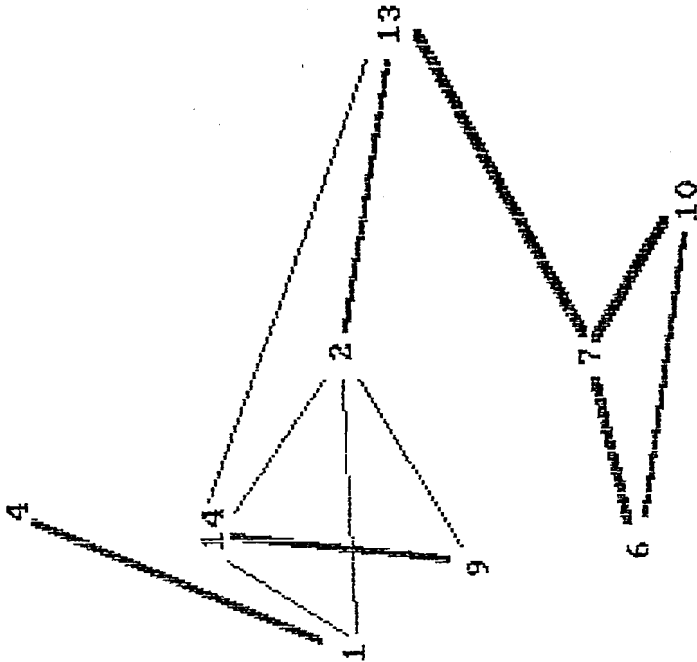
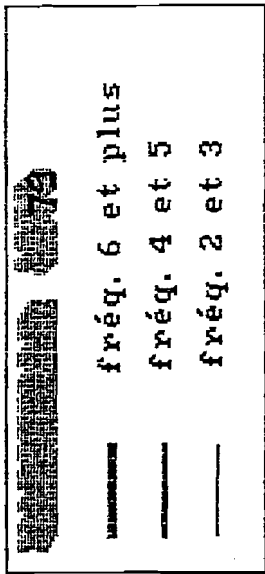
951 références analysées.

POLES DE RECHERCHE MARSEILLE 1987
 Pole principal 1 pharmacology






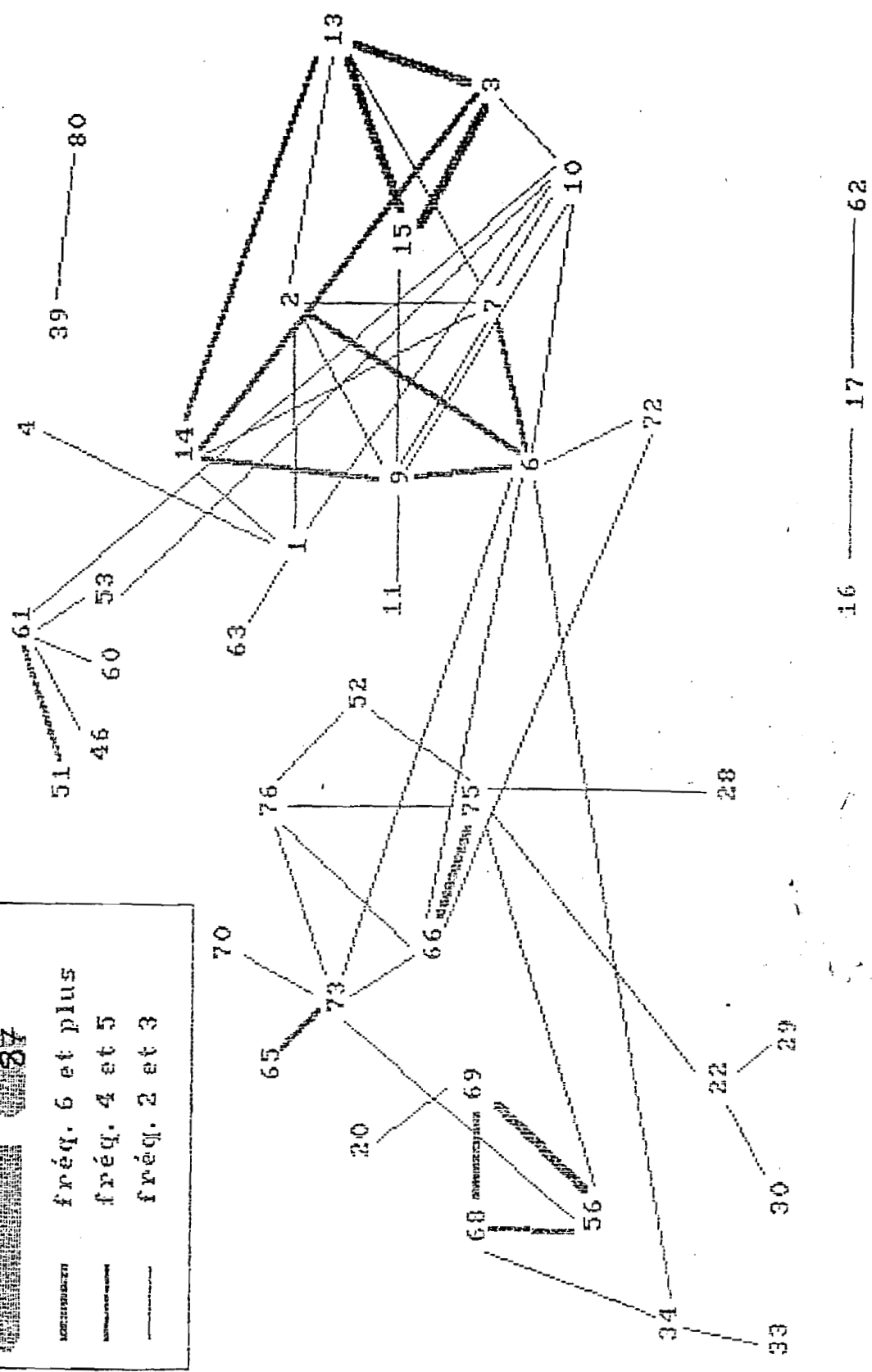
POLES DE RECHERCHE MARSEILLE 1979
 Pole principal 22 physical organic chemistry





87

	fréq. 6 et plus
	fréq. 4 et 5
	fréq. 2 et 3



ANNEXE: Codes des sections Chemical Abstracts: numéro et libellé

1	pharmacology
2	mammalian hormones
3	biochemical genetics
4	toxicology
5	agrochemical bioregulators
6	general biochemistry
7	enzymes
8	radiation biochemistry
9	biochemical methods
10	microbial biochemistry
11	plant biochemistry
12	nonmammalian biochemistry
13	mammalian biochemistry
14	mammalian pathological biochemistry
15	immunochemistry
16	fermentation and bioindustrial chemistry
17	food and feed chemistry
18	animal nutrition
19	fertilizers; solids; and plant nutrition
20	history; education; and documentation
21	general organic chemistry
22	physical organic chemistry
23	aliphatic compounds
24	alicyclic compounds
25	benzene; its derivate; and condensed benzoid compounds
26	biomolecules and their synthetic analogs
27	heterocyclic compounds(one hetero atom)
28	heterocyclic compounds(more than one hetero atom)
29	organometallic and organometalloidal compounds
30	terpenes and terpenoids
31	alkaloids
32	steroids
33	carbohydrates
34	amino acids; peptides; and proteins
35	chemistry of synthetic high polymers
36	physical properties of synthetic high polymers
37	plastics manufacture and processing
38	plastics fabrication ans uses
39	synthetic elastomers and natural rubber
40	textiles
41	dyes; organic pigments; fluorescent brighteners; and photographic sensitizers
42	coating; inks and related products
43	cellulose; lignin; paper; and other wood products
44	industrial carbohydrates
45	industrials organic chemicals; leathers; fats; and waxes
46	surface-active agents and detergents
47	apparatus and plant equipment
48	unit operations and proceses
49	industrial inorganic chemicals
50	propellants and explosives
51	fossil fuels; derivatives; and related products
52	electrochemical; radiational; and terminal energy technology
53	mineralogical and geological chemistry
54	extractive metallurgy
55	ferrous metals and alloys
56	nonferrous metals and alloys
57	ceramics
58	cement; concrete; and related building materials
59	air pollution and industrial hygiene
60	waste treatment and disposal
61	water
62	essential oils and cosmetics
63	pharmaceuticals
64	pharmaceutical analysis
65	general physical chemistry
66	surface chrmistry and colloids
67	catalysis; reaction kinetics; and inorganic reaction mechanisms
68	phase equilibriums; chemical equilibriums; and solutions
69	thermodynamics; thermochemistry; and terminal properties
70	nuclear phenomena
71	nuclear technology
72	electrochemistry
73	optical; electron; and mass spectroscopy and other related properties
74	radiation chemistry; photochemistry; and photographs and other reprographic processes
75	crystallography and liquid crystals
76	electric phenomena
77	magnetic phenomena
78	inorganic chemicals and reactions
79	inorganic analytical chemistry
80	organic analytical chemistry