

**DOCUMENTS DE TRAVAIL DE LA DSTI
1997/1**

**INDICATEURS BIBLIOMÉTRIQUES ET ANALYSE DES
SYSTÈMES DE RECHERCHE : MÉTHODES ET EXEMPLES**

Yoshiko Okubo

ORGANISATION DE COOPERATION ET DE DEVELOPPEMENT ECONOMIQUES

Paris

49700

Document complet disponible sur OLIS dans son format d'origine

Complete document available on OLIS in its original format

Documents de travail de la DSTI

La série de Documents de travail de la Direction de la science, de la technologie et de l'industrie a été créée dans le but de rendre accessible à un plus large public les analyses rédigées par des membres de la Direction ou par des consultants externes travaillant sur des projets pour l'OCDE. Les rapports sont d'une nature technique et/ou analytique et traitent de questions sur les données, la méthodologie et l'analyse empirique dans les domaines de travail de la Direction. Les Documents de travail sont en général disponible uniquement dans leur langue d'origine – anglais ou français – et présentent un bref résumé dans l'autre langue.

Des commentaires sur ces Documents seraient appréciés et pourront être adressés à l'OCDE, 2 rue André Pascal, 75775 Paris Cedex 16, France.

Les opinions exprimées sont celles des auteurs et ne reflètent pas nécessairement celles de l'OCDE ni celles des gouvernements des pays Membres.

Copyright OCDE, 1997

Les demandes de reproduction ou de traduction doivent être adressées à :
M. le Chef du Service des Publications, OCDE, 2 rue André Pascal, 75775 Paris, Cedex 16, France.

Indicateurs bibliométriques et analyse des systèmes de recherche : méthodes et exemples

Yoshiko Okubo*

Ce rapport qui s'apparente aux documents techniques des manuels de l'OCDE pour la mesure des activités de R-D ("famille Frascati") présente les principes de la bibliométrie et leurs applications à l'analyse des systèmes de recherche. La bibliométrie peut se définir comme la discipline qui mesure et analyse la production ("l'output") de la science sous forme d'articles, de publications, de citations, de brevets et autres indicateurs dérivé plus complexes. La bibliométrie constitue un instrument important dans l'évaluation des activités de recherche, des laboratoires et des chercheurs, ainsi que dans l'appréciation des spécialisations et des performances scientifiques des pays. Ce document, après avoir rappelé les conditions de développement de la bibliométrie, présente les bases de données sur lesquelles elle se fonde, puis les principaux indicateurs que l'on utilise. Vingt-cinq exemples commentés sont présentés à la fin du document illustrant les divers usages des méthodes bibliométriques. Ils permettent de mesurer la production scientifique appréciée par les volumes de publications, les impacts des publications sur les disciplines scientifiques déterminés par les co-citations, la coopération internationale matérialisée par les signatures conjointes, le contenu scientifique des brevets, etc.

This report – linked to the technical documents of the OECD manuals for the measurement of R&D activities ("Frascati family") – presents the essential elements of bibliometrics and its application to the analysis of research systems. Bibliometrics is based on the enumeration and statistical analysis of scientific output in the form of articles, publications, citations, patents and other more complex indicators. It is an important tool in evaluating research activities, laboratories and scientists as well as the scientific specialisations and performances of countries. The report, having set the background to the development of bibliometrics, presents the databases on which bibliometrics are elaborated as well as the principal indicators used. Twenty-five examples are presented at the end of the document, illustrating the various uses of bibliometrics methods for analysing research systems. These indicators measure scientific output by counting the number of publications, the impacts of publication on scientific disciplines by counting the number of co-citations, the extent of international co-operation as evidenced by joint signatures, the scientific content of patents, etc.

* Ce texte a été rédigé par Yoshiko Okubo, Laboratoire Stratégie & Technologie, École Centrale Paris, Grande voie des Vignes, 92295 Châtenay-Malabry Cedex. Téléphone : (33) 1 41 13 10 24. Fax : (33) 1 46 83 99 17. E-mail : yoshiko@cti.ecp.fr

L'auteur tient à remercier les personnes suivantes pour leur aide dans l'acquisition des données et leurs critiques : Jennifer Bond, Tibor Braun, L. Grauls, Shlomo Herskovic, Terttu Luukkonen, Aris Kaloudis, Maurits Pino, Rosa Sancho, Per O. Seglen, Gunnar Sivertsen, Henry Small, Anthony Van Raan, Jan C.G. Van Steen, et Gunnar Westholm, ainsi que Jean-Eric Aubert du Secrétariat de l'OCDE pour la mise en forme et la publication de ce texte.

TABLE DES MATIÈRES

Indicateurs bibliométriques et analyse des systèmes de recherche : méthodes et exemples.....	3
CHAPITRE 1. QU'EST-CE QUE LA BIBLIOMÉTRIE ?	10
CHAPITRE 2. L'AVÈNEMENT DE LA BIBLIOMÉTRIE	12
Historique	12
Le besoin d'évaluer la recherche.....	13
Bibliométrie et mesure de la science.....	14
CHAPITRE 3. LES BASES DE DONNÉES BIBLIOMÉTRIQUES	16
Les principales bases de données bibliométriques.....	16
Problèmes posés par les banques de données en général	17
Les types de "publications"	17
La base de données <i>Science Citation Index</i> (SCI)	18
La structure et le potentiel de la base SCI	18
La base <i>Science Citation Index</i> et ses limites	19
Pourquoi cite-t-on ?	19
Les "citations négatives"	20
Les "non cités"	20
Les "auto-citations"	20
L'impact linguistique	20
La ventilation par discipline scientifique.....	20
CHAPITRE 4. L'UTILISATION DES INDICATEURS BIBLIOMÉTRIQUES ET LES PRÉCAUTIONS A PRENDRE.....	22
Généralités	22
Le problème des co-signatures – "comptage entier" ou "comptage fractionné" ?.....	23
Le problème du champ de couverture des banques de données	24
La commercialisation des données et des analyses bibliométriques.....	25
Quel avenir pour les indicateurs bibliométriques ?.....	25
CHAPITRE 5. LES PRINCIPAUX INDICATEURS BIBLIOMÉTRIQUES ET LEURS APPLICATIONS.....	26
Introduction	26
Les indicateurs de l'activité scientifique et technologique.....	26
Le nombre des publications (exemples 1 à 4).....	26
Usages.....	26
Limites.....	27
Le nombre des citations (exemples 5 et 6)	27
Usages.....	27
Limites.....	27

Le nombre des co-signatures	27
Usages.....	28
Limites.....	28
Le nombre des brevets (exemples 10 et 11)	28
Usages.....	29
Limites.....	29
Le nombre de citations des brevets (exemple 12).....	29
Usages.....	30
Limites.....	30
Indicateurs relationnels (exemples 13 à 16)	30
Les co-publications	30
Usages.....	30
Limites.....	31
L'indice d'affinité (exemple 17)	31
Usages.....	31
Limites.....	31
Les liens scientifiques des citations (exemples 18 et 19).....	31
Usages.....	31
Limites.....	32
Corrélations entre publications scientifiques et brevets (exemples 20 à 22).....	32
Usages.....	32
Limites.....	32
Les co-citations (exemple 23).....	32
Usages.....	32
Limites.....	33
La "co-occurrence" des mots (exemple 24)	33
Usages.....	33
Limites.....	33
Techniques de visualisation des champs scientifiques et des pays (exemples 25, 25A et B)	33
Usages.....	33
Limites.....	34
LISTE DES EXEMPLES.....	35
RÉFÉRENCES	62
ANNEXE I. INDEXE DES CODES ISO	67
Exemple n° 1. Evolution de la production scientifique mondiale 1973-86 par principaux pays (%)	35
Exemple n° 2. Le nombre de publications par chercheur dans cinq pays, 1986	36
Exemple n° 3. Ventilation par pays-auteurs des publications sur le thème "génomique humaine", 1991 (%)..	37
Exemple n° 4. Spécialisation par discipline : le poids de la médecine clinique et de la physique dans les publications scientifiques nationales, 1981-86	38
Exemple n° 5. Pourcentage des publications ¹ citant et étant citées par d'autres publications ¹ , par principaux pays (moyennes 1984-86)	39

Exemple n° 6. Citations parues dans des publications en 1984, ventilées par principaux domaines scientifiques et par l'orientation de R-D en ingénierie et technologie	40
Exemple n° 7. Publications dans la science et l'ingénierie ayant un, deux, trois, quatre ou plus d'auteurs 1973-86 (%)	41
Exemple n° 8. Publications co-signées université-industrie aux États-Unis, par rapport à la totalité des articles signés des chercheurs industriels, par domaines scientifiques 1981-91 (%)	41
Exemple n° 9. Tendances des co-signatures internationales de quelques pays, 1981-86 (%)	43
Exemple n° 10. Ventilation des brevets déposés auprès des bureaux des brevets des États-Unis, de l'Europe et du Japon, par principaux pays (%)	44
Exemple n° 11. "Productivité" de la recherche pharmaceutique industrielle	45
Exemple n° 12. Citations des brevets américains par des brevets postérieurs, classés par pays inventeurs et secteurs propriétaires des brevets cités	45
Exemple n° 13. Le réseau international de la Ciba-Geigy Corporation, 1990	47
Exemple n° 14. Carte des liens de coopération internationale dans les sciences, 1981-85	48
Exemple n° 15. Position du Japon dans l'activité scientifique internationale de l'Allemagne, de la France et du Royaume-Uni en ingénierie (ENT) et sciences de la Terre et de l'espace (EAS), 1981-86	49
Exemple n° 16. Profils comparés – Strasbourg et France, 1986 Liens de coopération scientifique entre Strasbourg et 97 pays partenaires	50
Exemple n° 17. Indices d'affinité entre le Japon et la Suède 1981-86 (%)	51
Exemple n° 18. Répartition des citations dans le monde, 1984-88 (%)	52
Exemple n° 19. Citations entre revues spécialisées du domaine des polymères	53
Exemple n° 20. Relation à la science des brevets du monde entier déposés aux États-Unis	54
Exemple n° 21. Répartition par discipline scientifique des citations scientifiques dans les dépôts de brevets du monde entier aux États-Unis (moyennes, 1987-88)	55
Exemple n° 22. Production et utilisation de science en rapport avec les brevets (moyenne 1987-88) en poids mondial	56
Exemple n° 23. Agrégats ("clusters") de co-citations en 1992 et 1993 dans le domaine de la "chimie combinatoire"	57
Exemple n° 24. Carte du "champ de la technologie" dans l'optomécatronique	57
Exemple n° 25. Matrice des collaborations internationales, 1981-86 : 98 pays et huit domaines scientifiques	59

Exemple n° 25A. “L’arbre de longueur minimum” Similitudes des pays basées sur le profil de leurs coopérations internationales, 1981-86.....	60
Exemple n° 25 B. Positions relatives en coopération internationale des 98 pays déterminées par les huit champs scientifiques, 1981-86.....	60

INTRODUCTION

Naguère complètement ignorés par les décideurs politiques, les indicateurs de l'activité scientifique prennent place désormais au coeur des discussions sur les liens entre le progrès scientifique et technologique et le progrès économique et social. On a pris conscience de l'intérêt qu'il y a à fonder les opinions, et les choix qui en résultent, sur des évaluations quantitatives. L'examen des politiques scientifiques paraîtrait inconcevable aujourd'hui sans le recours à ces indicateurs de science et de technologie. Longtemps centré sur les indicateurs de ressources ("input"), comme les dépenses et le personnel de R-D, l'intérêt se porte maintenant de plus en plus sur les résultats ("output"), tout particulièrement ceux des activités technologiques (brevets, balance des paiements technologiques, commerce de haute technologie). Pour ce qui est de la science, ce sont les indicateurs bibliométriques qui s'imposent.

Dans ce document, nous esquissons d'abord une histoire de la bibliométrie, pour retracer les conditions de son émergence et définir son rôle dans l'évaluation de la science. Nous présentons divers indicateurs bibliométriques, en notant leur utilité et leurs limites, ainsi que des exemples d'applications pratiques. Cette étude a pour but de mettre en relief aussi bien les points forts que les faiblesses des indicateurs bibliométriques, et surtout les quelques précautions qui s'imposent. Ce survol ne prétend nullement être exhaustif.

La mesure bibliométrique est un outil qui permet d'observer l'état de la science et de la technologie à travers la masse des publications scientifiques, dans un contexte plus ou moins large. C'est un moyen de situer un pays dans le monde, une institution dans un pays, et même la place d'un scientifique dans sa communauté. Ces indicateurs scientifiques se prêtent – avec les précautions qui s'imposent – aussi bien à des analyses "macro" (par exemple, la part d'un pays donné dans la production mondiale des publications scientifiques pendant une période donnée) qu'à des études "micro" (par exemple, le rôle d'un institut dans la production de textes dans un domaine scientifique très précis). C'est une manière d'observer l'état de la science qui peut aider à mieux en comprendre la structure. En fournissant de nouvelles informations, la bibliométrie peut être une aide à la décision et à la gestion de la recherche. Elle ne peut pas, bien entendu, toute seule, justifier une décision, voire remplacer les experts. Les indicateurs bibliométriques sont des outils pratiques susceptibles d'être utilisés conjointement avec d'autres indicateurs.

Les produits de la science ne sont pas des objets mais des idées, des moyens de communication, des réactions aux idées des autres. On peut suivre à la fois les scientifiques et l'argent investi. Mais il est bien plus difficile de mesurer la science en tant que corpus d'idées ou les relations qu'elle entretient avec le système économique et social. Les indicateurs restent, pour l'instant, essentiellement une mesure fondée sur des observations sur la science et la technologie en tant que système d'activités, plutôt qu'un corpus de connaissances spécifique (National Science Foundation, 1989).

Chaque indicateur a ses avantages et ses limites et il faut éviter de les considérer comme des indices "absolus". Il faut tester la "convergence" (Martin and Irvine, 1985) des indicateurs pour relativiser les informations qu'ils fournissent.

La demande croissante émanant de la communauté scientifique elle-même pour les chiffres bibliométriques est un fait. La publication de rapports sur le modèle des “Indicateurs” de la National Science Foundation des États-Unis, qui s’est étendue à la plupart des pays industrialisés, témoigne de ce besoin. C’est qu’il s’agit de choisir bien, en connaissance de cause et avec l’aide de chiffres “objectifs”, des éléments utiles pour convaincre les membres d’un parlement, d’un conseil d’administration, ou les citoyens, voire les scientifiques et les chercheurs eux-mêmes.

CHAPITRE 1. QU'EST-CE QUE LA BIBLIOMÉTRIE ?

En 1969, Pritchard a proposé un mot nouveau, bibliométrie, pour un type de travaux qui existait depuis un demi-siècle. Le fait que Pritchard ait éprouvé le besoin de redéfinir un domaine couvert depuis 50 ans par l'expression *statistical bibliography* (Hulme, 1923) reflète l'apparition d'un nouveau domaine de recherches quantitatives. Selon Pritchard, la bibliométrie est : "l'application de méthodes mathématiques et statistiques aux livres et aux autres moyens de communication" (Pritchard, 1969, pp. 348-349).

La bibliométrie est devenue le terme générique pour toute une gamme de mesures et d'indicateurs spécifiques ; elle se donne pour objet de mesurer les productions ("output") de la recherche scientifique et technologique, à travers des données issues non seulement de la littérature scientifique mais aussi des brevets.

Les approches bibliométriques, qui permettent de décrire la science à travers ses résultats, reposent sur l'idée que l'essentiel de la recherche scientifique est la production de "connaissances" et que la littérature scientifique en est la manifestation constitutive. Par ailleurs, le brevet signifie un transfert des connaissances vers l'innovation industrielle et le passage à une valeur commerciale et sociale ; il fournit donc un indicateur pour mesurer le profit tangible d'un investissement intellectuel et économique.

Le principe qu'une publication (voir plus loin la discussion sur ce qui constitue une publication pour la bibliométrie) constitue la part essentielle de l'activité des scientifiques est, depuis des années, soutenu par les analystes de la science. Selon Price, est un scientifique "quiconque a, au moins une fois, publié un article scientifique" (Price, 1963). "Chaque fois qu'un homme travaille", écrit-il, "il produit des objets nouveaux et si ses résultats sont des publications, alors il a fait ce que j'appelle de la science" (Price, 1969). La formule "publier ou périr" ("*publish or perish*") indique que la publication des résultats de la recherche est au premier rang des activités des hommes de science.

Publier les résultats de leurs recherches est une obligation que les scientifiques doivent remplir (Merton, 1957b). Les connaissances nouvelles, mises au jour par les chercheurs, doivent être transformées en informations mises à la disposition de la communauté scientifique. Non seulement un scientifique doit mettre ses travaux à la disposition du grand public mais il est, à son tour, censé avoir accès à ceux de ses pairs. Ainsi la recherche s'élabore sur le terrain de "l'échange". Mais le fait que le système de publication scientifique s'est maintenu dans la science moderne est dû, paradoxalement, au désir des scientifiques de protéger leur propriété intellectuelle. Une nouvelle connaissance scientifique est la création personnelle d'un chercheur et la propriété attachée à sa découverte ne peut être assurée que par la publication (Merton, 1957a).

Le "système des récompenses" (*reward system*) fondé sur la reconnaissance du travail ne fait que souligner l'importance de la publication ; la seule manière de faire connaître les résultats d'une recherche dans le monde se fait à travers une publication.

Ainsi, la publication a trois objectifs : la diffusion des résultats scientifiques, la protection de la propriété intellectuelle, l'acquisition de la notoriété.

Les scientifiques doivent publier leurs travaux et la publication justifie leur existence. "Une publication érudite", ajoute Price, "n'est pas un paquet d'informations mais une expression de l'état d'un chercheur (ou d'un groupe de chercheurs) à une époque particulière. On ne publie pas, contrairement à une certaine croyance, un fait, une théorie, ou une découverte, mais un mélange des trois. Un article scientifique est, à la fois et à des degrés divers, un concept, des données et une hypothèse. Si l'article est l'expression de plusieurs personnes travaillant sur un même sujet, on peut deviner, à partir de l'article lui-même, le type de relations existant entre ces personnes" (Price, 1963).

Aujourd'hui, la bibliométrie est appliquée à des champs très divers¹ :

- l'*histoire des sciences*, où elle éclaire la recherche sur l'évolution des disciplines scientifiques, en suivant les mouvements historiques révélés à travers les résultats obtenus par les chercheurs ;
- la sphère des *sciences sociales*, où elle soutient, grâce à l'examen de la littérature scientifique, l'analyse de la communauté scientifique, de sa structure dans une société donnée, des motivations et des réseaux de chercheurs ;
- la *documentation*, où elle permet de recenser le nombre de revues (*journals*) par bibliothèque, ou encore d'identifier les revues qui forment le cœur, la partie secondaire et la périphérie d'une discipline (analyse de la quantité de revues qu'il faut associer pour couvrir 50 pour cent, 80 pour cent ou 90 pour cent des informations dans un domaine scientifique donné) ;
- la *politique de la science*, où elle donne des indicateurs pour mesurer la productivité et la qualité scientifique, fournissant ainsi des éléments d'évaluation et d'orientation des efforts de R-D.

Les techniques bibliométriques ont évolué dans le temps et continuent de le faire, en ce qui concerne le comptage des publications attribuées à un pays, à une institution ou à un auteur ; le comptage des citations, pour mesurer l'impact des travaux publiés sur la communauté scientifique ; et le comptage des co-citations (nombre de fois où deux articles sont cités ensemble dans un même article), etc., afin de mieux détailler les informations et de disposer d'indicateurs plus "performants". Les résultats sont présentés sous diverses formes, y compris sous forme de "cartes" qui permettent de visualiser les relations entre les acteurs et d'élargir la gamme des instruments d'analyse.

CHAPITRE 2. L'AVÈNEMENT DE LA BIBLIOMÉTRIE

Historique

L'idée d'examiner les publications remonte au début de ce siècle. En 1917, Cole et Eales ont publié une analyse statistique de l'histoire de l'anatomie comparée. Cette effort fait date dans l'histoire de l'analyse bibliométrique, car ces auteurs sont parmi les premiers à avoir utilisé la littérature publiée pour construire un profil quantitatif des avancées dans un champ de recherche. L'article décrit les apports de la technique bibliométrique et les problèmes qu'elle pose, dont certains ne sont pas encore résolus.

Un autre travail, utilisant cette fois les brevets, est celui de Hulme (1923). En mettant en corrélation les brevets et la littérature scientifique afin de mesurer le progrès social de la Grande-Bretagne, Hulme a été l'initiateur d'une méthodologie moderne de l'histoire des sciences.

Ensuite, Lotka (1926) a montré les fréquences de distribution de la production scientifique. Il a été sans doute l'un des premiers à associer la notion de productivité au dénombrement, en utilisant l'index décennal des *Chemical Abstracts* et des *Geschichtstafeln der Physik* d'Auerbach. Il a également proposé une mesure qualitative des travaux scientifiques sur la base des données qui permettent de sélectionner les contributions les plus éminentes. Lotka a constaté que le nombre d'articles publiés n'est pas distribué de manière homogène, et que la productivité a tendance à se concentrer sur un nombre limité de chercheurs.

En 1935, Cunningham a publié une étude sur la littérature biomédicale, et en 1952, Boig et Howerton en ont publié une sur la publication en chimie. Cependant, jusque dans les années 60, les travaux publiés dans le domaine qui allait devenir la bibliométrie étaient rares. Le fait que l'expression "*statistical bibliography*" ait été utilisée dans la littérature moins de cinq fois entre 1923 et 1962 montre combien cette activité demeurait confidentielle (Pritchard, 1969).

Les années 70 voient littéralement exploser le nombre d'études bibliométriques : elles marquent une deuxième époque de la bibliométrie, dont le point de départ est l'apparition d'une base de données reposant sur les citations des articles scientifiques, le *Science Citation Index* (SCI), qui a été fondé en 1963 par Eugene Garfield aux États-Unis et qui a ouvert la voie à tous ceux qui tentent de mesurer la science par des méthodes quantitatives et objectives.

Au départ, Garfield voulait créer un moyen de permettre aux chercheurs de trouver, rapidement et efficacement, des articles publiés dans leur domaine de recherche (Garfield, 1968). Mais il a bientôt élargi ses études à l'évaluation des références recueillies : "Lorsque l'entreprise scientifique devient de plus en plus large et complexe, et son rôle dans la société de plus en plus critique, il devient aussi de plus en plus difficile, coûteux et nécessaire d'évaluer et d'identifier les individus et les groupes dont les contributions sont les plus importantes" (Garfield, 1979b). Ainsi, Garfield cherchait à faire de l'analyse de la citation un outil légitime et pratique pour évaluer la production scientifique.

L'existence du SCI a non seulement déclenché de nombreuses études bibliométriques, mais elle a favorisé la naissance d'une nouvelle génération de bibliométriciens qui revendiquaient cette approche

comme étant la “science de la science” (Price, 1965). Partisan écouté de cette méthodologie, lui-même physicien de formation, Price a tenté de porter sur la science un regard indépendant de celui des scientifiques : selon lui, la science peut être mesurée par la publication, et la science peut être analysée en dehors des scientifiques. Les scientifiques, toujours selon Price, sont des spécialistes qui, en dehors de leur propre domaine de recherche, ne sont plus spécialistes. Il écrit : “A l’instar des analyses économiques, devenues d’une aide très précieuse dans la prise de décisions gouvernementales et industrielles, et à côté de leur intérêt proprement académique, il se pourrait que nous soyons en train d’assister à la naissance d’une évaluation scientifique comme d’une analyse du monde de la science” (Price, 1964). Price pressentait que, dans un avenir proche, l’analyse des citations allait être utilisée comme un complément de l’évaluation par les pairs.

Dans ce domaine, des travaux russes qui remontent aux années 30 associent les analyses scientifiques et les sciences sociales ; ces études avaient pour objet la description méthodologique des diverses disciplines. Les systèmes de mesures mis au point ont permis de créer un domaine nouveau, la *Naukometrika* (littéralement, “la mesure de la science”), précurseur de la bibliométrie.

La revue internationale *Scientometrics*² s’est donné pour objectif de publier des articles sur tous les aspects quantitatifs de la “science de la science” ; elle présente une part importante des méthodes et des études bibliométriques contemporaines et constitue un forum très actif de discussions, parfois vives, entre représentants des différentes écoles de bibliométriciens.

La première génération de bibliométriciens a élaboré des concepts et des mesures techniques qui ont été par la suite affinés à des fins d’évaluation des activités scientifiques. Mais d’après Wade (1975), un usage pratique en est fait dès avant 1975, par exemple pour l’évaluation des politiques des “conseils de recherche”, l’analyse de la recherche universitaire, l’appréciation des besoins de création de nouveaux instituts de recherche dans des domaines en émergence.

Le besoin d’évaluer la recherche

Les sociétés industrialisées ont été très favorables au développement de la science (Bush, 1960). Dès leur début en 1957, les enquêtes de la NSF montraient que le public américain croyait que la science et la technologie contribuaient grandement au progrès de la société (National Science Foundation, 1989, pp. 170-172).

Stimulé par la compétition avec l’Union soviétique, les États-Unis font un effort considérable de R-D dans les années 60, effort accompagné par la création d’organismes et d’institutions. L’Europe, l’Union soviétique et le Japon ont ensuite suivi le même chemin.

Un changement s’est produit dans les années 70 : la science n’est plus alors une aventure dans laquelle la société peut investir infiniment et généreusement. La première phase de ce changement a été provoquée par le ralentissement de la croissance économique, mais elle était également le fait d’une attitude plus critique qui prenait en compte les effets négatifs de la recherche scientifique : la science et la technologie coûtent cher, mais les investissements dans la recherche ne permettent pas forcément de résoudre les problèmes de l’environnement ou les problèmes sociaux, par exemple ceux dus à l’écart entre les pays industrialisés et le Tiers monde.

Ainsi, on a été conduit à se préoccuper de la rentabilité de la recherche fondamentale en particulier et les chercheurs étaient de plus en plus perçus comme des “producteurs de science” qui doivent rendre compte des fonds qu’ils reçoivent.

Par ailleurs, du fait des révoltes étudiantes, l'image des universités s'est ternie, tout comme celle de l'autorité des scientifiques et des diplômés ; ces événements ont éveillé la méfiance des citoyens à l'égard de la science et de la technologie. Il fallait dorénavant viser la production d'une valeur ajoutée qui épargne les ressources naturelles et crée moins de pollution, établir un système de recherche plus efficace tirant mieux parti des stocks d'intelligence. C'est dans ce contexte que l'évaluation de la recherche scientifique s'est trouvée mise en valeur.

On a ainsi eu recours aux méthodes des sciences sociales et humaines (sciences "molles") pour l'analyse des sciences "dures" ; il a fallu disposer de critères et de mesures quantitatifs, c'est-à-dire trouver des méthodes de quantification, constituer des indicateurs et trouver les moyens de les comparer. L'établissement d'une mesure de la (ou des) science(s) est devenu inévitable. Ce mouvement a ouvert le champ de l'analyse de la science et de la technologie et favorisé l'entrée des bibliométriciens sur la scène de la politique scientifique.

Bibliométrie et mesure de la science

Partout, les pouvoirs publics ont progressivement senti le besoin d'analyses critiques de leur politique scientifique et technologique. Certains ont jugé suffisant de créer des services administratifs au sein de leurs administrations de la recherche (ministères de l'enseignement, de l'industrie, etc.). D'autres ont préféré la formation de spécialistes et la mise au point d'indicateurs dans un cadre universitaire qui encourage la confrontation des idées.

Aux États-Unis, la NSF a publié le rapport intitulé *Science and Engineering Indicators* pour la première fois en 1972. Les responsables présentaient ainsi leur travail : "L'objet final de notre travail est d'établir un jeu d'indices qui permette de révéler les points forts comme les faiblesses de la science et technologie américaines, en termes de capacité et de performance, par rapport aux objectifs nationaux. Si de tels indicateurs sont développés dans les années à venir, leur rôle sera d'aider à une meilleure répartition ainsi qu'à une meilleure gestion des ressources financières en science et technologie, et de diriger la recherche et le développement à l'avantage de notre société" (NSF, 1972). Dans les rapports qui ont suivi, le rôle des indicateurs bibliométriques s'est considérablement accru.

A partir de cette époque, la bibliométrie est de plus en plus orientée vers la politique de la science. Des groupes de bibliométriciens de différentes écoles ont proposé divers moyens de mesurer la croissance de la science, en liant leurs méthodologies aux évaluations. C'est ainsi que la bibliométrie commence la tâche délicate qui consiste à chercher à contribuer à l'évaluation. Bon nombre de ces méthodes ont été présentées lors de séminaires spécialisés à travers le monde.

En un demi-siècle d'histoire, la bibliométrie a ainsi gagné sa place comme un des instruments de mesure de la science, aussi bien dans les pays occidentaux industrialisés que dans les pays de l'Est et les pays en voie d'industrialisation, tel que l'Inde. En 1989, l'OCDE a consacré un chapitre du supplément au Manuel de Frascati au secteur de l'enseignement supérieur, confirmant ainsi la place de la bibliométrie dans l'analyse de la science (OCDE, 1989, pp. 49-53).

Les Pays-Bas et le Royaume-Uni ont été parmi les premiers à publier régulièrement des études sur la science en utilisant des indicateurs bibliométriques. Des groupes de recherche de ces pays ont été des pionniers dans l'élaboration théorique et l'exploitation pratique dans ce domaine (Irvine et Martin, 1980 ; Martin et Irvine, 1984 ; Leven, 1982 ; Moed *et al.*, 1983), en introduisant notamment, pendant les années 80, des mesures bibliométriques dans les analyses de la politique scientifique. En 1987, le ministère de l'Éducation, de la Science et de la Culture du Japon (Monbusho) a confié à une équipe de

bibliométriciens la réalisation d'une étude comparative du nombre d'articles scientifiques publiés dans sept grands pays afin "d'établir un indicateur pour apprécier l'activité de la recherche du Japon dans le contexte international" (Ministère de l'Éducation, de la Science et de la Culture, 1987).

Cependant, dans tous les pays, il a fallu beaucoup de temps pour que l'approche bibliométrique soit acceptée comme une mesure de la science, tant dans les milieux politiques que chez les scientifiques. Certains scientifiques continuent de manifester leur hostilité : nombreux sont ceux qui connaissent peu cette méthodologie et qui n'aiment pas être analysés (et encore moins évalués) par une mesure quantitative du niveau d'activité de la recherche, "mesure grossière, qui n'est qu'un simple comptage des publications" (Rappa, 1989, p. 28). Il n'est pas aisé de passer - soit psychologiquement soit intellectuellement - d'une mesure à l'échelle des pays ou des disciplines scientifiques à celle d'une évaluation au niveau de chaque chercheur : "C'est un scandale intolérable" (Chauvin, 1991, p. 782). Certains scientifiques ont proposé une méthode d'évaluation qui permettrait de juger les candidats non pas sur une liste de publications tirée d'une base de données plus ou moins anonyme, mais sur quelques articles qu'ils considèrent eux-mêmes comme les plus représentatifs de leurs travaux. Les chercheurs estiment que c'est uniquement à travers une évaluation par les pairs (*peer review*) qu'ils doivent être jugés. Cette vue est partagée par une grande partie de la communauté scientifique actuelle (Ourisson, 1991).

Au niveau de l'individu, la bibliométrie mesure la productivité en matière de recherche, mais elle ne met pas nécessairement en évidence les qualités pédagogiques ni la compétence pour l'enseignement. Les réactions des scientifiques sont bien naturelles et soulignent le besoin d'interaction entre ceux qui sont évalués et ceux qui évaluent.

Dans le processus d'évaluation, le dialogue entre "créateurs de la science" et "analystes bibliométriciens" peut être constructif. Il peut contribuer à modifier les données et les méthodes utilisées, mais surtout l'interprétation des résultats. Nombre d'experts de "bibliométrie évaluative" confirment que la discussion des résultats doit toujours faire partie de l'évaluation des chercheurs (Moed *et al.*, 1983). Dans le dialogue entre analystes et analysés, la vérification des chiffres est essentielle. Il est clair que, pour une évaluation individuelle, l'utilisation des indicateurs bibliométriques demande bien plus de vigilance que pour une description générale de la science au niveau d'un pays.

Avec le temps, ces méthodes sont devenues de plus en plus reconnues, mais peut-être plus au sein des communautés non scientifiques qu'au sein de la communauté scientifique elle-même (Chelimsky, 1991). Bon nombre de pays publient des statistiques sur le modèle des *Science and Engineering Indicators* de la NSF : *Australian Science and Technology at a Glance 1990* (Australie), *Science Indicators Compendium 1991* (Canada), *Science and Technology Indicators 1991* (Japon), *Science et Technologie - Indicateurs 1992* (France), *S&T Indicators Report 1994* (Pays-Bas), *Politique Scientifique et Technologie - Bilan et perspectives* (OCDE), et le rapport européen sur les indicateurs scientifiques et technologiques.

Outre *Scientometrics*, les revues qui publient des articles utilisant les méthodes bibliométriques sont de plus en plus nombreuses : citons, entre autres, *Research Policy*, *Science and Public Policy*, *Research Evaluation*, *Journal of the American Society for Information Science*, *Rapport de l'Observatoire des Sciences et des Techniques* (France).

CHAPITRE 3. LES BASES DE DONNÉES BIBLIOMÉTRIQUES

Les principales bases de données bibliométriques

La source de la bibliométrie est toujours une base de données. Diverses bases, mises en place par des sociétés commerciales ou par des institutions publiques ou privées, sont utilisées pour illustrer les résultats de l'activité scientifique et technologique (en données brutes). Soumises à un traitement spécifique, elles peuvent être utilisées pour établir des indicateurs bibliométriques. La plupart de ces bases sont spécialisées, quelques-unes seulement sont généralistes. Parmi les bases les plus utilisées (de plus amples détails figurent à l'annexe), on peut citer :

- *Chemical Abstracts* : Base de données spécialisée dans les sciences physico-chimiques. Elle est produite par la société américaine, Chemical Abstracts Service, auprès de l'American Chemical Society ; elle enregistre en moyenne environ 500 000 références par an provenant d'environ 10 000 revues.
- *Compendex* : Base de données spécialisée dans l'ingénierie et la technologie. Elle est produite par la société américaine Engineering Information ; elle enregistre en moyenne environ 150 000 références par an provenant d'environ 4 500 revues scientifiques.
- *Embase* : Base de données spécialisée dans les sciences médicales. Elle est produite par la société néerlandaise, Excerpta Medica ; elle enregistre en moyenne environ 250 000 références par an provenant d'environ 3 500 revues.
- *Inspec* : Base de données spécialisée dans les sciences physiques. Elle est produite par l'Institute of Electrical Engineers en Angleterre ; elle enregistre en moyenne environ 200 000 références par an provenant d'environ 2 200 revues.
- *Pascal* : Base de données généraliste sur divers domaines. Elle est produite par l'Institut de l'Information Scientifique et Technique du Centre National de la Recherche Scientifique (CNRS) en France ; elle enregistre en moyenne environ 450 000 références par an provenant d'environ 6 000 revues.
- *Science Citation Index* : Base de données pluridisciplinaire produite par une société américaine, l'Institute for Information Science (voir plus loin).

En ce qui concerne les données sur les brevets, les bases WPI(L) de la société anglaise Derwent Inc. et la banque de données de Computer Horizon Inc. (CHI) sont les plus souvent utilisées. Les premières se classent parmi les bases pluridisciplinaires et internationales, répertoriant les demandes de brevets publiées et les brevets délivrés par 30 offices nationaux, tandis que la banque de données de CHI est surtout fondée sur les statistiques de l'Office des brevets des États-Unis.³

Ces bases sont généralement disponibles en ligne et/ou sur CD-ROM.

Problèmes posés par les banques de données en général

Le choix d'une banque de données pour l'élaboration d'indicateurs bibliométriques est directement lié aux buts recherchés et aux questions auxquelles la base doit fournir la réponse. Chaque base de données a son contenu et ses règles de recensement ; aucune n'est identique à aucune autre. Ainsi, sur un sujet précis, la quantité d'articles (ou autres unités de mesure) variera selon la banque choisie. La qualité (le détail des ventilations par discipline scientifique, par exemple) différera aussi, notamment en fonction de la sélection des revues dont les données sont tirées. Pour une même étude, cette diversité peut entraîner des résultats divergents, d'où la nécessité de chercher des données aussi cohérentes que possible. En effet, lorsque les résultats divergent selon les sources choisies, il n'existe aucun moyen objectif de distinguer celle (ou celles) qui reflète(nt) le mieux la réalité de la production scientifique. Chaque utilisateur est donc obligé de commencer son étude par le choix de la base de données la plus appropriée à ses besoins précis ; pour cela, il doit tenir compte, au préalable, des atouts, des faiblesses et des limites des différentes bases.

Pour les études bibliométriques au niveau "macro", les bases choisies doivent être représentatives, mais ne doivent pas nécessairement couvrir l'intégralité des données. La combinaison ou l'intégration dans l'analyse d'informations provenant de sources (banques) différentes minimise toujours le risque de ne pas être représentatif ou exhaustif.

Par ailleurs, les fiches bibliographiques qui constituent les diverses bases de données ont au départ été établies pour fournir des informations surtout qualitatives, et non pas en vue d'un comptage quelconque de publications ou d'articles. C'est pourquoi ces données, après extraction, ont besoin d'un traitement spécifique qui les rende utilisables pour la production d'indicateurs. A cette fin, plusieurs modes de traitement ont été élaborés par les bibliométriciens et employés dans leurs analyses.

Lorsque l'on veut étudier l'évolution de la science dans une certaine discipline, on est obligé de proposer ou d'entreprendre l'agrégation du champ scientifique puisque les banques de données ne sont pas automatiquement classées selon la spécialité souhaitée. Si l'on souhaite analyser le domaine des "céramiques" par exemple, l'analyste peut interroger la base de données : *i*) par les titres des revues spécialisées dans ce domaine (identifiés par l'analyste lui-même) ; *ii*) par les mots-clés associés à sa recherche (sélectionnés aussi par l'analyste lui-même) ; ou *iii*) par une liste sélective des revues utilisées par les spécialistes. Selon le mode d'interrogation, les résultats varient. Attribuer une recherche à un champ ou une discipline spécifique est une tâche délicate, surtout parce qu'un programme de recherche peut très bien associer diverses disciplines.

Les types de "publications"

De même, il incombe à l'utilisateur des bases de données de décider, en fonction de l'étude menée, ce qu'il faut inclure ou exclure dans le comptage. Il existe plusieurs types de publications dans une banque de données typique : articles, notes, synthèses, lettres au directeur de la publication, comptes rendus, notices, discussions, livres, etc. Pour des raisons évidentes, l'article est le mode d'expression de base de la transmission des connaissances nouvelles. Mais pour presque tous les autres types de publications, le choix appartient aux analystes. Ceux-ci doivent choisir quels contenus des bases utilisées seront incorporés dans l'étude bibliométrique. A ce propos, on a constaté, non sans ironie, que "l'on peut facilement créer une courbe qui traduise la baisse, la hausse ou la stabilité de la science . . . selon le choix des types de publication" (Leydesdorff, 1991).

La méthodologie du dénombrement des publications, des techniques de traitement des données et des concepts de base relève donc des choix des utilisateurs. Par conséquent, l'utilisation des indicateurs bibliométriques exige une grande prudence et il est nécessaire de comparer des résultats obtenus à partir de plusieurs bases de données, notamment si l'on étudie des disciplines scientifiques assez pointues ou récentes, donc encore peu structurées.

La base de données *Science Citation Index* (SCI)

La structure et le potentiel de la base SCI

Pour mesurer la quantité d'un stock scientifique, on doit utiliser des données homogènes. Il est nécessaire que la base de données repose sur des critères définis et mesurables, afin que les analystes puissent préciser la communauté examinée. C'est dans ce contexte que la base de données *Science Citation Index* (SCI), créée par l'Institute for Scientific Information aux États-Unis, prend son importance.

Son "inventeur", E. Garfield, a considéré le coût-avantage comme le critère fondamental pour définir le champ couvert par une base de données. "Il est . . . impossible de tenir compte de toutes les revues scientifiques existantes, car personne ne connaît la quantité de revues publiées, puisqu'il n'existe pas de définition de ce qui constitue une revue" (Garfield, 1972).

Garfield a d'abord calculé le nombre de revues scientifiques nécessaires pour optimiser l'obtention d'un maximum d'informations scientifiques. Il a adopté une loi développée par des scientifiques de l'information, elle-même fondée sur le travail de Bradford (1950), qui permet de sélectionner le nombre de revues nécessaires pour obtenir un maximum d'informations. Elle montre qu'il faut examiner entre 500 et 1 000 revues pour obtenir 95 pour cent de la littérature "significative" publiée dans un champ donné.

Garfield a, par la suite, associé la loi de dispersion de Bradford à une loi de concentration établie par lui-même (Garfield, 1972). La loi de Bradford définit un champ scientifique, mais si l'on veut en couvrir plusieurs dans une base de données, on peut se demander s'il faut multiplier le nombre de revues pour un champ par le nombre de champs. Selon Garfield, comme une partie importante des disciplines se recoupe, le coeur de la littérature pour l'ensemble de ces disciplines peut également être couvert par environ 500 à 1 000 revues.

Il fallait d'abord mettre au point une méthode qui permît d'identifier ces 500 à 1 000 revues. Pour cela, Garfield a utilisé "le nombre de citations" comme un des critères d'une recherche "significative" ; en utilisant les citations, les auteurs se réfèrent aux travaux précédents pour soutenir, décrire ou développer un point spécifique de leur travail. L'acte de citer une publication scientifique est ainsi une indication de l'importance que la communauté accorde à ce travail de recherche et la citation peut être considérée comme un critère pour sélectionner les revues scientifiques les plus estimées à partir des publications qu'ils contiennent.

Premièrement, on compte le nombre de fois qu'un article est cité dans une revue donnée. Ensuite, on calcule le facteur d'impact (*impact factor*) en divisant le nombre de citations par le nombre d'articles contenus dans cette revue. Cela permet d'éliminer tout avantage lié à aux dimensions de la revue, et rend la citation proportionnelle au nombre d'articles.

La base de données SCI prend ainsi en compte les revues scientifiques les plus utilisées, les plus reconnues, voire les plus influentes dans le monde, cela étant mesuré par leurs "indices de citations". Elle limite le champ couvert aux revues scientifiques à rayonnement mondial, qui représentent le "coeur" de la

production scientifique dans des domaines précis ; elle élimine en quelque sorte les recherches qui ne sont pas présentées dans le “club des meilleurs”. Ainsi, le monde du “courant majeur” (*mainstream*) est limité à un groupe spécifique de revues. Le problème pour une base de données est de définir une frontière entre “puissants” et “faibles”. Est-ce la qualité, la quantité, ou d’autres principes de mesure qui font la différence ? De toute manière, pour un chercheur, c’est la reconnaissance internationale de sa production qui le situe sur l’échelle des valeurs. Vers 1980, la SCI tenait compte de quelque 3 100 revues scientifiques ; à l’époque, il existait quelque 70 000 revues scientifiques dans le monde ! C’est la raison pour laquelle on a baptisé le monde du SCI “le courant majeur” (*mainstream*) (Frame, 1977). La sélection des revues couvertes par le SCI introduit ainsi un aspect qualitatif dans le comptage des publications.

Mis à part sa position presque monopolistique, le fait que le SCI est actuellement utilisé de manière aussi intensive n’est pas à mettre uniquement au compte de l’emploi des citations. Cette base couvre un vaste domaine de la science ; elle est pluridisciplinaire, alors que la plupart des autres bases sont plus spécialisées. Avec le SCI, on peut faire une étude panoramique de la science, grâce à son traitement homogène des sciences exactes et naturelles ; mais ISI a aussi élargi sa collection des bases de données et l’on peut maintenant obtenir de bonnes informations sur les sciences sociales et humaines à partir des *Social Science Citation Index* (SSCI) et *Arts & Humanities Citation Index* (AHCI).

Le SCI enregistre les affiliations (adresses, appartenances institutionnelles, etc.) de tous les auteurs de chaque article, tandis que la plupart des autres bases n’enregistrent que le premier auteur des articles co-signés (ce qui, parfois, constitue une “place d’honneur” pour des chercheurs renommés qui n’ont pas nécessairement beaucoup contribué aux travaux). A présent, le *SCI* et *Physic Briefs* sont parmi les rares bases de données qui aient adopté une politique d’affiliations multiples. Cette méthode dispose d’atouts de poids, surtout à une époque où des études sur l’internationalisation de la science sont bien avancées. C’est grâce à l’enregistrement de toutes les affiliations des co-signatures qu’un programme informatique peut sélectionner un article écrit par des chercheurs dans des laboratoires et des pays différents. Cela donne au SCI une position dans la recherche bibliométrique que les bibliométriciens ne peuvent ignorer.

La base Science Citation Index et ses limites

Pourquoi cite-t-on ?

Cependant, le SCI a aussi des limites qui sont communes à beaucoup de bases bibliométriques. Les raisons qui conduisent l’auteur d’un article scientifique à citer d’autres publications sont complexes. Depuis des années, les sociologues de la science ont analysé la signification de la citation ; le fait de se référer aux travaux des autres n’est pas toujours lié à l’originalité ou à l’importance d’un travail, voire à sa “qualité”.

La citation peut dépendre aussi de la capacité à diffuser un article auprès d’un large public. Un scientifique célèbre dirige souvent de nombreux élèves et ses articles auront plus de chances d’être cités que ceux de scientifiques moins influents (surtout si le nom du “patron” figure-- à tort ou à raison – parmi les co-auteurs). Le poids de la structure sociale de la discipline n’a pas nécessairement un lien direct avec la qualité de la recherche. On peut se référer aux “grands” maîtres pour leur rendre hommage (ou par simple tactique), plutôt qu’en reconnaissance d’un travail que l’on admire (*The Economist*, 18 janvier 1992, p. 87). Le travail d’un chercheur possédant une technique ou une méthodologie expérimentale, pas forcément de grande qualité mais utile, sera cité chaque fois que ses méthodes sont utilisées.

La citation est une mesure de l’impact général de l’influence d’un article, ou de l’influence de ses auteurs, sur la communauté scientifique ; c’est un paramètre socio-épistémologique complexe qui induit

probablement un facteur de qualité, mais ce facteur n'est ni équivalent, ni corrélé d'une manière non équivoque, à la qualité scientifique (Seglen, 1992).

Les "citations négatives"

Une référence peut être aussi de nature négative. Un auteur peut être cité pour une recherche sujette à polémique, voire pour une erreur méthodologique. Ainsi, encore une fois, la citation ne mesure pas toujours la valeur du travail mais plutôt l'impact d'une recherche ou d'un scientifique.

Les "non cités"

En outre, la masse de scientifiques cités est fortement limitée. Plus de la moitié (55 pour cent) des articles publiés dans les revues scientifiques recensées dans le SCI n'ont pas été cités une seule fois dans les cinq ans qui ont suivi leur parution. Le taux de non cités varie selon les disciplines : dans la catégorie des sciences de l'ingénieur, il s'agit de plus de 72 pour cent des publications (Pendlebury, 1991).

Les "auto-citations"

Le problème des auto-citations, c'est-à-dire les références faites par un chercheur à ses travaux antérieurs, est brièvement examiné plus loin.

L'impact linguistique

Un autre point qui a été souvent évoqué, critiqué et parfois analysé, est le fait que cette base favorise nettement les scientifiques anglophones (Otsu, 1983 ; Kobayashi, 1987). D'après Garfield, "il n'est ni facile ni économique d'inclure les revues qui n'utilisent pas l'alphabet romain dans la base de données du SCI" (Garfield, 1975). Cette observation montre que, pour être accepté dans le "club des meilleurs", il faut écrire dans une langue majeure : la langue anglaise est actuellement nettement accentuée et le système est auto-entretenu (Garfield, 1988). Les chercheurs des pays non anglophones qui publient en anglais ont d'ailleurs une présence comparativement plus grande : c'est le cas, par exemple, pour les pays scandinaves (Sivertsen, 1991).

La ventilation par discipline scientifique

L'un des principaux atouts des banques de données bibliométriques est la possibilité d'utiliser des classifications par discipline scientifique bien plus détaillées et désagrégées que celles qui servent habituellement aux enquêtes pour mesurer les dépenses et le personnel de R-D. Habituellement, les citations se réfèrent à des articles, etc., au sein d'une même catégorie ou sous-catégorie d'un domaine ou d'une discipline.

Or, une difficulté apparaît pour l'utilisation du SCI lorsque l'on étudie l'état de la science à partir d'une base pluridisciplinaire qui couvre un vaste domaine et qui ne peut prendre en compte les différentes dynamiques et les spécificités des citations d'une discipline à l'autre. *A priori*, les critères de sélection sont les mêmes pour chacune des revues couvertes dans la base. Cependant, les traditions et les habitudes de publication et de citation varient selon les champs et cela joue un rôle dans la représentativité des données. Par exemple, la plupart de la littérature mondiale en physique, chimie et biomédecine (publiée et citée) est

bien représentée dans SCI. En revanche, il existe des problèmes de couverture pour les géosciences, la recherche en biologie sur le terrain, l'ingénierie et la technologie et les mathématiques, et, dans une certaine mesure, la médecine clinique. Les causes sont à rechercher dans le fait que les communications dans certaines disciplines sont concentrées dans les principales revues internationales spécialisées et aussi dans le fait que certaines revues ont une "influence" moins étendue. Le problème de la langue est aussi plus grand dans certains domaines scientifiques.

Tout cela vaut aussi pour la pratique de la citation ; pour préciser le poids d'un pays dans les citations, par exemple, il faut tenir compte des spécificités de chaque domaine qui influent sur l'indice de citations. Par exemple, il est avéré qu'à court terme les articles de biomédecine sont plus cités en moyenne que les articles de mathématiques et de médecine clinique. Ces corrélations doivent être prises en compte, surtout lors de l'interprétation des chiffres. Des corrections devraient être introduites, mais elles exigeraient une meilleure compréhension de la structure des domaines scientifiques analysés. Ceci oblige à une certaine prudence dans l'interprétation des citations ; une coopération entre les praticiens d'une discipline et les bibliométriciens semble particulièrement nécessaire dans ce cas.

CHAPITRE 4. L'UTILISATION DES INDICATEURS BIBLIOMÉTRIQUES ET LES PRÉCAUTIONS A PRENDRE

Généralités

Les analyses bibliométriques utilisent de nombreux paramètres : la publication scientifique (articles, etc.), la co-signature, le brevet, la citation, la co-citation, les mots-clés associés (*co-words*) ; ce sont des mesures indirectes de la communauté scientifique, de sa structure et de sa production. On trouvera des exemples de l'utilisation de ces indicateurs au chapitre 5.

Les données et analyses bibliométriques fournissent des informations sur l'orientation ou le dynamisme scientifiques d'un pays (ou d'une autre unité) et sur sa participation aux activités mondiales en science et en technologie, autrement dit, sur la présence de cette unité dans la communauté nationale et internationale. L'analyse des coopérations permet de dessiner les réseaux scientifiques tissés, en mettant en relief les relations établies entre pays, institutions ou chercheurs ou l'impact des grands programmes [CERN (Laboratoire européen de la physique des particules), Organisation mondiale de la santé (OMS), etc.]. La bibliométrie permet aussi d'analyser la structure des disciplines scientifiques et leurs relations. Les données et indicateurs bibliométriques peuvent servir comme outils ou, tout au moins, comme une aide pour décrire et exprimer les questions qui se posent dans le monde de la science.

Comme pour d'autres domaines, il est important de noter que les indicateurs obtenus à travers les bases bibliométriques doivent être relativisés. Les indicateurs sont fondés sur une logique de la comparaison : les chiffres absolus ne sont pas, par eux-mêmes, indicatifs, ils ne prennent leur valeur qu'en comparaison avec d'autres ensembles.

Dans les analyses, il est recommandé de prendre en compte un aussi grand nombre de données que possible, ce qui permet une compensation statistique des distorsions pouvant affecter chacune des petites entités prises individuellement.

Les limitations des données utilisées en bibliométrie tiennent essentiellement aux divers modes de communication employés par les scientifiques pour se transmettre des informations en dehors des circuits habituels des revues scientifiques. Les communications entre scientifiques par voie orale, notamment, ne sont pas capturées par les statistiques, tout comme les rapports internes communiqués entre universités, laboratoires, groupes de recherche ou ceux des pays associés dans des comités, programmes ou laboratoires communs. Manquent également d'importantes monographies, et plus encore les échanges électroniques entre chercheurs, qui se développent rapidement. Ainsi, les communications couvertes par les méthodes bibliométriques "traditionnelles" sont des échanges "formalisés" ; les communications informelles ne sont pas prises en compte et ne le seront vraisemblablement jamais.

De même, l'approche traditionnelle est encore plus limitée pour tout ce qui touche à la recherche industrielle et à celle liée à la défense. La communication entre science (universitaire surtout) et industrie ne se fait qu'avec de grands retards, en raison du fait que l'industrie souhaite protéger ses découvertes, notamment avant le dépôt de demandes de brevets, et publie ses résultats le plus souvent sous une forme

réduite. Les articles publiés par un laboratoire industriel donnent volontairement une vue limitée des buts de la recherche, qui vise en général la création de produits et de procédés qui sont sujets à la compétition commerciale.

Quant à la recherche liée à la défense (souvent associée à la recherche industrielle), une grande partie ne fait jamais partie de la communication scientifique habituelle ; pourtant, fort importante au niveau de la technologie, elle comprend souvent aussi les travaux les plus avancés en recherche fondamentale.

Le problème des co-signatures – “comptage entier” ou “comptage fractionné” ?

Une des ambiguïtés de la méthode bibliométrique se trouve dans la diversité des modes de comptage. Par exemple, la classification des publications “co-signées”, c’est-à-dire les articles écrits par plus d’un auteur, soulève depuis longtemps un débat majeur parmi les bibliométriciens (Martin, 1991 ; Braun *et al.*, 1991 ; Leydesdorff, 1991 ; Kealey, 1991). Comment peut-on mesurer la participation des auteurs à des travaux scientifiques lors d’une coopération ? Peut-on accorder “un crédit entier” à chaque individu pour sa part ? Ou bien a-t-il droit à “un dixième” d’un article écrit avec neuf autres scientifiques ? Un pays a-t-il tissé “un lien entier” avec un autre pays en coopération internationale ou bien crée-t-il “le tiers d’un lien” dans une collaboration à laquelle participent trois pays ?

En pratique, quand un article est co-signé par des chercheurs de différents pays, les bibliométriciens ont, en gros, deux façons d’attribuer un crédit aux pays concernés :

- Certains attribuent un crédit complet, c’est-à-dire comptent “1” pour chacun des pays auteurs de co-signatures (“comptage entier” ou *whole counting*).
- D’autres divisent une co-signature par le nombre de pays d’origine des auteurs et attribuent un crédit fractionné à chacun des pays (“comptage fractionnaire” ou *fraction counting*). Cette méthode repose sur “la logique mathématique” : afin d’atteindre les 100 pour cent, le crédit de chaque pays doit être partagé et “1” doit être divisé par le nombre de pays co-signataires pour chaque co-publication internationale.

Chaque manière de compter possède sans doute sa propre logique, mais encore faut-il que le fractionnement soit totalement maîtrisé par les bibliométriciens et compris par les utilisateurs.

Pour les utilisateurs scientifiques et politiques des indicateurs de la science, le comptage entier est bien plus compréhensible et facile à traduire. Dix pour cent de crédit international accordé à un pays indique que dix articles sur les 100 articles internationaux publiés comptent au moins un auteur de ce pays. En revanche, dans le cas du comptage fractionnaire, il est difficile d’expliquer la signification des 10 pour cent, qui peuvent être la somme de dix publications ou plus. (Braun *et al.*, 1991).

Plus important encore, le comptage fractionnaire accorde une valeur inférieure à une co-signature internationale qu’à la signature d’un article national, lors du comptage de la performance nationale. Et plus un article a de partenaires internationaux, moins chaque pays concerné aura de crédits. Pourquoi un pays serait-il mieux crédité, dans les statistiques bibliométriques, pour une publication d’auteurs nationaux que pour un article co-signé dans le cadre d’une coopération internationale ? (Leydesdorff, 1991). C’est précisément pour ces raisons que certains bibliométriciens considèrent le comptage par fractions une procédure inférieure ; surtout lorsque les données représentent une quantité importante, ils sont d’avis que “le comptage entier est dans la plupart des cas la meilleure solution” (Van Raan et Tijssen, 1990).

Il est difficile de choisir entre les deux méthodes. Pourtant, tant que les deux seront employées, les analyses de la science et la technologie au niveau d'un pays ou d'un laboratoire peuvent varier et donner parfois des résultats contradictoires. Les distorsions que ces comptages accumulent en faveur ou au détriment de certains pays risquent de rendre problématiques les comparaisons internationales.

Le problème du champ de couverture des banques de données

D'autres débats concernent un point fondamental de la mesure bibliométrique : comment peut-on mesurer la tendance de la publication scientifique sur quelques années à travers des bases de données qui, elles, évoluent d'une année sur l'autre ? Ce problème, lié à l'histoire du développement de la bibliométrie, apparaît rétrospectivement comme le résultat des efforts faits pour rendre plus représentatifs les outils bibliométriques. Certains revues scientifiques disparaissent, d'autres changent de nom ou fusionnent et, surtout, de nouveaux titres paraissent. L'ISI suit cette évolution et actualise régulièrement la liste de revues retenues pour le SCI, ce qui crée une rotation annuelle d'environ 7 pour cent (Garfield, 1979a). En 1964, l'ISI avait incorporé environ 610 revues dans le SCI ; en 1981, le nombre a atteint 3 600 ! On assiste parallèlement à une augmentation du nombre des articles, qui est passé de 100 000 en 1964 à 500 000 en 1981 (ensemble des revues mobiles – *non-fixed journal set*) (Institute for Scientific Information, 1981).

Cependant, cette augmentation peut représenter une gêne pour certains, quand il s'agit de mesurer les tendances de la performance scientifique nationale dans la durée. Pour le comptage des publications, il peut sembler préférable que le nombre et la composition des revues dans une base soient stables pendant la période étudiée, pour que la mesure de l'état de la science d'un pays donné soit comparable d'une année sur l'autre. La croissance du nombre de références d'un pays pourrait être le simple résultat de l'addition de nouvelles revues et non pas une mesure de véritable croissance de la productivité. Les données doivent être traitées sous une forme qui les rende comparables d'une année sur l'autre (Anderson *et al.*, 1989). Afin de rendre plus aisée l'interprétation des données, on a proposé de traiter un nombre constant de revues (*fixed journal set*) dans l'ensemble des périodiques constitutifs de la base, soit environ 2 100 revues parmi celles que l'ISI a dépouillées en 1973. Ainsi, la liste des revues est demeurée "stable" entre 1973 et 1980. La liste des revues scientifiques choisies a de nouveau été révisée en 1981, afin de tenir compte de l'évolution de la science. Les nouveaux titres incorporés dans le SCI ne sont pas ajoutés au "*fixed journal set*" pendant les périodes "gelées" (1973-1980 et 1981-1986) ; en revanche, les revues qui ont été supprimés du SCI disparaissent de la base.

Limiter les références à un nombre donné de revues conduirait à ignorer le dynamisme naturel des connaissances, et ce "gel" risquerait de créer un monde "artificiel" de la science.⁵ De nouveaux domaines de recherche, tels que la supraconductivité ou les recherches sur le SIDA, qui apparaissent dans des revues nouvelles consacrées à ces domaines, risquent de ne pas pouvoir être étudiés. Les performances des pays industrialisés sont mieux représentées dans l'ensemble non constant que dans l'ensemble "constant", du fait que les scientifiques de ces pays ont tendance à choisir de nouveaux titres plutôt que les revues classiques pour présenter leurs travaux. C'est précisément dans les nouvelles revues que les nouveaux domaines de la recherche font leur apparition (Kealey, 1991). C'est ainsi que "la représentation de la recherche (fournie par la base constante) est sensiblement conservatrice et statique"(Callon et Leydesdorff, 1987).

Il y a donc, ici aussi, deux modes de traitement des données SCI, pour mesurer la performance de la science : *i*) à partir de "revues non constantes", la collection dont la quantité varie ; et *ii*) à partir de "revues constantes", la collection qui diminue avec les années. Seuls les points de départ (1973 et 1981) sont fixes ; ensuite, les résultats n'évoluent pas de la même façon. Aucune des séries n'est donc "stable" ou "constante". La question cruciale est de savoir lequel des sous-ensembles est le plus représentatif ou

produit les indicateurs les plus fiables pour la mesure des tendances des performances nationales en science (Martin, 1991). Dans certaines spécialités, les discordances sont telles que les tendances divergent avec le temps (The Royal Society, 1986).

Il n'est pas difficile de chiffrer les publications, mais quand il s'agit de donner un sens à ces chiffres, l'affaire devient plus complexe. Les chiffres ne "parlent" pas d'eux-mêmes, il faut les interpréter en tenant compte des distorsions réelles et artificielles des données et de la méthode utilisée pour le dénombrement.

La commercialisation des données et des analyses bibliométriques

Une autre difficulté que rencontrent les bibliométriciens est directement liée aux pressions commerciales qui sont apparues dans ce domaine.

Ce type d'études est recherché : les données et l'analyse bibliométrique ont aujourd'hui un intérêt commercial et il n'est pas rare que des contrats pour des travaux (extractions, traitements et analyses des données) portent sur des montants forts élevés. Chacune des diverses écoles bibliométriques a développé sa méthode et de ce fait les résultats peuvent varier. La commercialisation a augmenté la compétition entre elles, mais, en même temps, la pression du marché a encouragé la professionnalisation du milieu.

A l'heure actuelle, les débats méthodologiques agitent les bibliométriciens. Les méthodes sont souvent proches, mais il n'y a ni normes (Glanzel et Schoepflin, 1994) ni consensus sur les meilleures méthodes ou applications de la bibliométrie.

Quel avenir pour les indicateurs bibliométriques ?

Longtemps réservée aux évaluations et analyses des recherches académiques et des "grands programmes" publics (nationaux et internationaux), la bibliométrie est maintenant en train de gagner du terrain dans d'autres secteurs, surtout grâce à la mise au point d'indicateurs pour décrire l'évolution des différents types de coopération (Nederlands Observatorium van Wetenschap en Technologie, 1994 ; et Katz et Hicks, 1996).

Les indicateurs de base de la bibliométrie doivent encore beaucoup progresser. Les bibliométriciens poursuivent leurs efforts pour appliquer et améliorer les indicateurs existants. Un des champs d'étude les plus intéressants concerne la prospective scientifique et technologique, domaine pour lequel des indicateurs ont été mis au point (Leydesdorff, 1995 ; Noyon et Van Raan, 1995). Des travaux sont également en cours pour développer des "indicateurs intégrés", c'est-à-dire pour associer plusieurs indicateurs pour représenter des activités scientifiques et technologiques (Niwa et Romizawa, 1995). A côté de cette approche macroscopique, des essais au niveau "micro" tentent de représenter les mouvements de la science vers la technologie, à travers l'évolution de réseaux de groupes de recherche (Hirasawa, 1995). Par ailleurs, on cherche à élargir les méthodes d'aide à l'évaluation en se fondant sur la qualité des revues scientifiques (Magri et Solari, 1996). Ces efforts devraient être particulièrement utiles pour l'analyse des politiques scientifiques.

Chaque indicateur a ses avantages et ses limites. Il faut prendre garde de ne pas les considérer comme des indices "absolus" ; ils sont complémentaires. Les différentes procédures et méthodes bibliométriques doivent être utilisées en parallèle, malgré des résultats parfois contradictoires, aussi longtemps qu'elles apportent des informations utiles et respectent les normes scientifiques et professionnelles. Malgré ses limites, la bibliométrie fournit une mesure quantitative globalement objective de la production scientifique.

CHAPITRE 5. LES PRINCIPAUX INDICATEURS BIBLIOMÉTRIQUES ET LEURS APPLICATIONS

Introduction

Ce chapitre présente les indicateurs bibliométriques les plus connus, avec des exemples pratiques tirés de publications internationales et nationales. Les exemples ont été choisis plus pour leur caractère exemplaire que pour l'actualité de leur contenu. Ils sont assortis d'un bref commentaire et d'observations méthodologiques. Une bonne partie de la littérature technique sur la bibliométrie est en langue anglaise ; c'est pourquoi la version originale de certains exemples a été gardée dans ce chapitre.

La première partie est consacrée aux principaux indicateurs quantitatifs des activités de la science et de la technologie, tandis que la deuxième partie traite surtout des indicateurs dits "relationnels", c'est-à-dire des indicateurs mesurant les liens et les interactions entre différents acteurs des systèmes de S-T, surtout du point de vue international. Il y a également une brève présentation des techniques de "visualisation" des variables scientifiques par des méthodes d'analyses multidimensionnelles.

Les indicateurs de l'activité scientifique et technologique

Le nombre des publications (exemples 1 à 4)

Cet indicateur illustre la production scientifique, mesurée par le comptage de "publications", terme ici utilisé pour signifier différents supports de textes scientifiques (livres, revues, journaux, comptes rendus, rapports, articles, etc.).

Usages

Le comptage des publications fournit une première mesure, simplifiée et approximative, de la quantité de travail produit par un scientifique, un laboratoire, une faculté, une équipe nationale et/ou internationale de R-D, un pays, etc. Le nombre absolu de ces publications constitue, en lui-même, un indicateur bibliométrique grossier, mais c'est seulement en examinant ces données de base par rapport à d'autres "masses" que l'on obtient des mesures plus significatives de la puissance *relative* des sujets étudiés. Ainsi, dans un domaine ou une discipline précis, le dynamisme de la recherche d'un pays, d'une équipe, etc., peut être suivi et son évolution dans le temps mesurée. Si, par la suite, les données de base sont divisées par le nombre des chercheurs ou par les sommes d'argent investies, on obtient des indicateurs "dérivés" pouvant permettre, en quelque sorte, une analyse de la "productivité" des travaux en question.

Limites

Il est raisonnable d'utiliser le nombre de publications comme indicateur quand on se trouve face à des chiffres élevés : un pays, une université, un laboratoire, un domaine de recherche, etc., sont d'autant mieux représentés que les chiffres sont importants. Mesurer la production d'un chercheur individuel, uniquement en utilisant ce type d'indicateur, est moins à conseiller, étant donné qu'une telle mesure ne tient pas compte de la qualité des publications, même si on prend en considération le degré de sélectivité des revues incluses dans la base. Les publications peuvent aussi représenter un temps de recherche, un volume et une qualité de travaux fort différents. Ces indicateurs ne permettent pas de mesurer la quantité ou la qualité du travail que représente un article, et lorsque l'article est co-signé par plusieurs personnes, le rôle de chacune n'est connu que des auteurs eux-mêmes.

Le nombre des citations (exemples 5 et 6)

On peut considérer les citations comme une mesure de "l'impact" des articles cités et aussi de leur opportunité et de leur intérêt. Il est présumé qu'une publication doit avoir une certaine qualité pour avoir un impact auprès de la communauté scientifique

Usages

Les auteurs se citent les uns les autres pour des raisons variables. En gros, les citations peuvent être divisées en deux groupes: Dans l'un, la citation des travaux antérieurs permet de mettre en évidence les nouveautés contenues dans l'article ; dans l'autre, l'auteur reconnaît, fait un hommage aux travaux précédents.

Les données sur les citations fournissent également des indications sur la manière dont certains domaines de la science et de la technologie sont liés (les co-citations sont discutées plus loin).

Limites

Un auteur a tendance à citer les travaux de sa propre communauté scientifique et les oeuvres d'auteurs "à la mode". Les raisons qui conduisent un auteur à choisir entre des travaux importants, à citer l'un plutôt que l'autre, sont nombreuses et impossibles à identifier. Ces travaux n'ont pas nécessairement influencé sa recherche.

Un chercheur peut également citer ses propres travaux, augmentant ainsi le nombre de citations qui lui seront créditées. Le phénomène des "auto-citations" existe bien et justifie les critiques sur la fiabilité de cette mesure. Le problème se pose toutefois relativement peu lorsque les chiffres analysés sont élevés.

Le nombre des co-signatures

Le nombre des co-signatures (ou coauteurs) d'une publication est un indicateur qui mesure la coopération nationale ou internationale (l'internationalisation).

Usages

L'analyse des co-signatures vise à identifier les coopérations à travers des publications signées par au moins deux chercheurs. On peut mesurer la quantité des travaux exécutés par plusieurs scientifiques à l'échelle individuelle ou institutionnelle, ainsi qu'au niveau national ou international. C'est un paramètre pour mesurer la croissance (ou le déclin) des recherches en coopération par rapport à une recherche faite par un seul scientifique. L'analyse chronologique des co-signatures est une mesure de la pénétration de la coopération internationale dans la "production" de la science et de la technologie nationales.

Limites

En principe, l'affiliation (l'adresse) retenue est celle du lieu de travail du (ou des) chercheur(s) et non pas celle de leur domicile ou pays d'origine. L'identification et l'enregistrement d'une publication est fonction de la manière dont l'affiliation est donnée, ce qui peut poser problème. C'est notamment le cas des affiliations institutionnelles ; certains instituts ou laboratoires dépendant, par exemple, de très grandes universités ou organismes de recherche (type CNRS en France) peuvent figurer dans les banques de données sous des dénominations différentes.

La difficulté de dénombrer les co-signatures réside aussi dans la diversité du calcul (le problème du "compte fractionnaire" et du "compte entier" examiné au chapitre 4). Selon l'approche choisie, les résultats varient. Il s'agit aussi bien du traitement de plusieurs coauteurs (de pays différents) d'un seul article, que du traitement d'un seul auteur ayant plusieurs affiliations, par exemple un scientifique temporairement détaché de son laboratoire d'origine pour faire de la recherche dans une institution étrangère.

Dans ce dernier cas, si l'auteur ne désigne que son laboratoire d'hôte, sa publication est entièrement créditée à ce laboratoire. En revanche, s'il indique ses deux affiliations, elle est créditée aux deux institutions (et pays) et prend ainsi l'allure d'une co-signature internationale. Ces dérives sont toutefois peu sensibles à grande échelle.

Le nombre des brevets (exemples 10 et 11)

Les statistiques sur les brevets fournissent des éléments pour mesurer les résultats des ressources investies dans les activités de recherche et développement, et tout particulièrement l'évolution des changements techniques dans le temps. Les brevets constituent, au départ, un moyen légal de protéger les inventions mises au point par des firmes, des institutions ou des personnes physiques et peuvent, à ce titre, être considérés comme des indicateurs d'inventivité. Les statistiques sur les brevets servent de plus en plus d'indicateurs de science et de technologie ; on trouve ainsi dans les documents des brevets nombre d'éléments qui se prêtent à des analyses bibliométriques.

En 1994, l'OCDE a mis en diffusion générale un manuel, "Les données sur les brevets d'invention et leur utilisation comme indicateurs de la science et de la technologie - Manuel Brevet 1994" [OCDE/GD(94)114]. L'Organisation recueille aussi dans ses banques de données de S-T un certain nombre de séries sur les brevets.

Dans les statistiques de l'OCDE (voir le "Manuel Brevet 1994" et *Principaux indicateurs de la science et de la technologie* – PIST), les données portent sur les demandes de brevets déposées (et non sur les brevets délivrés ou accordés) selon les procédures nationales, européenne et internationales.

Quatre types de données de brevets sont concernées :

- le nombre de “demandes résidentes” qui émanent d’inventeurs résidant dans le pays concerné pendant une période donnée : cet indicateur donne une idée de la production de technologie ; on peut en tirer des informations supplémentaires comme, par exemple, le “coefficient d’inventivité” (= demandes résidentes/10 000 habitants) ;
- le nombre de “demandes non-résidents” de brevets, déposées par les inventeurs ne résidant pas dans le pays : cet indicateur reflète la pénétration technologique ;
- le nombre de demandes nationales de brevets qui est la somme des demandes résidentes et non-résidents : il indique, en quelque sorte, la taille du “marché technologique” que représente ce pays ;
- le nombre de demandes de brevets déposées à l’étranger par les inventeurs résidant dans le pays concerné : cet indicateur reflète la “diffusion technologique” de ce pays.

Usages

On peut utiliser le comptage des brevets pour identifier la place d’une invention et le rôle de chaque inventeur dans la mise au point de nouvelles techniques ; c’est donc une mesure de l’activité novatrice et de la capacité technologique à l’échelle des nations et des branches et entreprises industrielles. Les premiers travaux d’utilisation des statistiques sur les brevets, en tant qu’indicateurs de la S-T, ont porté sur des objets bien identifiés, des molécules par exemple. Par la suite, on a mesuré des technologies en compétition les unes avec les autres, ainsi que le niveau d’invention de pays en compétition autour d’une invention majeure.

Ce sont les brevets américains qui ont d’abord été utilisés comme indicateurs “d’output”. Les brevets déposés aux États-Unis sont soumis à une analyse approfondie du caractère original de l’invention revendiquée. Un avantage de ces documents est qu’ils contiennent des descriptions et références très détaillées sur l’invention et fournissent des informations importantes (comme des citations) d’intérêt bibliométrique.

Limites

La propension des industriels à breveter leurs inventions varie selon la branche industrielle et d’une entreprise à l’autre ; un certain nombre d’améliorations technologiques majeures n’aboutissent pas à des brevets. De même, la “qualité” des brevets n’est pas systématiquement du même niveau ; les brevets n’ont pas la même signification en termes d’innovation technique et d’avenir économique. Il n’est donc pas recommandé de comparer des dépôts de brevets pour diverses technologies ou différents secteurs industriels. Cependant, dans un domaine macroscopique bien déterminé, celui des pays par exemple, on peut entreprendre des comparaisons. En dépit des limites qu’ils présentent, les brevets sont et seront de plus en plus utiles comme source d’informations pour une mesure approximative de l’innovation.

Le nombre de citations des brevets (exemple 12)

Cet indicateur mesure l’impact de la technologie (plus peut-être que l’impact de la science).

Usages

Il n'y a pas encore de méthode communément acceptée pour mesurer les brevets en termes de valeur absolue ou relative, mais on peut utiliser des citations de brevets comme indicateur de S-T. La première page d'un brevet contient généralement une référence aux brevets qui ont déjà été acceptés sur le même sujet. L'examineur du brevet propose ces références dans le cours de la procédure d'examen.

La citation des brevets est une façon de présenter "l'état de l'art", c'est-à-dire ce qui a déjà été réalisé dans les domaines associés à l'invention proposée en relation à la nouveauté et la signification de l'invention présentée. Ces brevets ont des chances d'être significatifs dans le domaine considéré. C'est pourquoi les citations de brevets précédemment obtenus sur un sujet proche de celui du brevet déposé peuvent servir d'indicateur pour mesurer l'importance du brevet cité pour la technologie que l'on désire protéger. Il arrive aussi que le déposant fournisse lui-même des citations de brevets dans son document de dépôt. Celles-ci sont pourtant moins utilisées dans les analyses bibliométriques.

Limites

Les citations choisies par les examinateurs soulèvent des questions sur les raisons pour lesquelles ceux-ci citent des références différentes de celles des déposants eux-mêmes. Les examinateurs ne sont pas des spécialistes et les citations peuvent refléter davantage leur poids juridique que leur caractère novateur. Par ailleurs, les citations proposées par les déposants ne sont pas encore acceptées comme une véritable mesure significative de l'importance du brevet cité, puisque celui-ci pourrait avoir été choisi pour des raisons autres que son importance scientifique. Il faut prendre conscience des limites de ces mesures, car les brevets peuvent être écrits pour cacher des inventions majeures derrière des avancées mineures afin de déjouer les concurrents. Les entreprises, aidées en cela par leurs conseillers juridiques, présentent des différences considérables dans leur façon de protéger leurs travaux de recherche.

Indicateurs relationnels (exemples 13 à 16)

Les co-publications

Cet indicateur mesure les interactions et les relations scientifiques entre réseaux, équipes, institutions et pays.

Usages

Une co-publication est le résultat d'une coopération entre les représentants de chaque entité et pays participant à un programme de recherche défini et exécuté en commun. Cette recherche crée des "liens" entre les acteurs (scientifiques, laboratoires, institutions, pays, etc.) ayant produit ensemble une publication scientifique. Le nombre total des liens instaurés par des participants spécifiques peut être défini, visualisé et mesuré au moyen des co-signatures. A partir d'un indicateur de co-signature, on peut ainsi dessiner ces relations. Sur ce principe, on peut construire une matrice avec, dans chacune des cellules, le nombre de co-signatures entre un auteur (ou des auteurs) de la ligne et un (ou plusieurs) de la colonne. Cet indicateur permet d'identifier les principaux partenaires des opérations de recherche et de décrire les réseaux scientifiques.

Limites

L'utilité de ces indicateurs est directement liée à la façon dont les questions d'affiliation et du comptage des coauteurs ont été traitées (problèmes abordés dans le chapitre 4).

L'indice d'affinité (exemple 17)

L'indicateur utilisé pour évaluer le taux relatif des échanges scientifiques d'un pays donné (A) avec un autre pays (B), pendant une période de temps donnée (et éventuellement dans un domaine scientifique précis), et ceci par rapport à l'ensemble de la coopération internationale de ces deux mêmes pays pendant la même période, est appelé "indice d'affinité".

Cet indicateur donne une "vue double" de ces liens, qui peuvent se mesurer, par exemple, en termes d'articles co-signés ; il s'applique, bien entendu, à d'autres entités que des pays (entreprises, groupes géographiques, etc.).

La formule du calcul de l'indice d'affinité (développé par le Laboratoire d'évaluation et de prospective internationales (LEPI) du CNRS (France) est la suivante :

$$\frac{COP(A - B)}{COP(A - MD)} \times 100$$

où COP(A - B) représente le nombre de liens scientifiques (coopérations) entre A et B et COP(A - MD) représente le nombre des liens de coopération entre A et le monde.

Usages

Cet indicateur mesure non seulement les liens entre pays mais aussi le niveau d'équilibre entre eux, autrement dit les "rapports de force" qui caractérisent ces échanges ; il montre aussi bien les domaines les plus forts que les plus faibles. En examinant l'évolution des indices d'affinité en fonction du temps, on obtient une indication des changements intervenus dans les rapports scientifiques bilatéraux.

Limites

L'indice d'affinité n'est applicable qu'à condition qu'il y ait une certaine "masse" de liens de coopération pendant la période, et ceci dans les deux sens. Il est préférable d'utiliser cet indicateur pour des collaborations scientifiques entre deux parties qui se rapprochent en termes de masse scientifique.

Les liens scientifiques des citations (exemples 18 et 19)

Cet indicateur mesure des rapports d'influence entre communautés scientifiques.

Usages

A travers les citations, on peut dessiner les réseaux d'influence entre différentes communautés scientifiques. Ces interactions mettent en évidence les évaluations conduites par les pairs sur les travaux scientifiques passés et en cours.

Limites

Un certain nombre des problèmes relatifs à cette approche ont été abordés dans le chapitre 4.

Corrélations entre publications scientifiques et brevets (exemples 20 à 22)

Cet indicateur illustre les liens (interactions) entre les sciences (mesurées par les publications) et les technologies (reflétées par les brevets).

Usages

Beaucoup d'informations peuvent être extraites des brevets et des documents qui les accompagnent, par exemple, des références aux publications scientifiques, qui sont parfois reprises dans les banques de données spécialisées. La relation entre les connaissances scientifiques (publications) et les technologies qui les utilisent peut être analysée par des références ou des citations faites par les inventeurs et/ou par les examinateurs des brevets déposés.

Deux types d'indicateurs sont proposés : le premier associe la science et la technologie à travers les citations scientifiques et les citations des brevets ; le deuxième mesure la période qui s'écoule entre la publication des articles scientifiques et le dépôt des brevets (décalage temporel).

L'indicateur d'intensité (ou de "proximité scientifique") d'une activité industrielle ou technologique est construit à partir du nombre relatif des citations de publications scientifiques dans les brevets déposés par ces secteurs.

Limites

Les corrélations entre brevets et publications n'ont pas encore fait l'objet d'analyses systématiques et les avis restent partagés sur leur importance et leur signification. En effet, le rôle du brevet est d'abord juridique ; le fait que ses auteurs veulent à la fois démontrer leur liens technologiques et cacher l'essentiel du contenu rend problématique l'utilisation statistique et analytique de ces données.

Les co-citations (exemple 23)

Les co-citations mesurent le nombre de fois que deux publications sont citées simultanément dans un même article. Cet indicateur illustre des réseaux thématiques et l'influence (l'impact) des auteurs. La méthode des co-citations représente, dans sa finalité, les réactions de la communauté scientifique aux résultats de la recherche.

Usages

Les agrégats ("*clusters*") des co-citations fournissent une description des sujets de recherche similaires et proches et décrivent les travaux complémentaires dans la spécialité étudiée, elle-même mesurée par les citations. On peut ainsi identifier et dessiner sur une "carte" les communautés de chercheurs intégrées dans un réseau particulier. Ces agrégats permettent aussi de montrer comment des champs et des sous-champs évoluent à travers le temps.

Limites

En ne décrivant qu'une partie de la construction des connaissances, les co-citations fournissent une analyse de la science fort sélective qui se réfère bien plus à la littérature scientifique qu'à la littérature technologique.

La “co-occurrence” des mots (exemple 24)

L'indicateur précédent (co-citations) examine le nombre de fois que deux publications sont citées ensemble. Celui-ci examine le nombre de fois (la fréquence) que deux mots (“*co-words*”), dans un champ particulier de la S-T, ont été utilisés ensemble dans des publications ou des brevets. Pour chaque mot, la co-occurrence avec un autre mot, ainsi que la fréquence, est analysée. Ces mots sont propres à chaque thème de recherche et sont choisis par des experts en la matière.

Cette méthode prend pour hypothèse qu'à partir de ces mots, il est possible d'identifier et de visualiser des réseaux spécifiques d'un type donné de recherche, en vue d'en étudier l'évolution. Dans les publications scientifiques et les brevets, la présence de ces mots traduit une similitude de concepts intellectuels chez les chercheurs. Ils représentent ainsi des signaux, indiquent des associations qui peuvent être représentés sous forme de graphiques lexicaux (“*leximapps*”). La fréquence des mots associés est utilisée pour construire une carte (diagramme stratégique) qui représente les thèmes majeurs du domaine étudié et leurs relations.

Usages

Cette méthode a été utilisée, par exemple, dans la description du rôle d'une agence gouvernementale pour consolider et transformer un réseau de chimie macromoléculaire.

Limites

Cette méthode soulève des problèmes pour l'interprétation des résultats. Les mots ne sont pas séparables de leur contexte syntaxique et il n'y a pas, semble-t-il, de manière systématique d'interpréter la carte. Les utilisateurs soulignent l'importance des analyses au “niveau micro”.

Techniques de visualisation des champs scientifiques et des pays (exemples 25, 25A et B)

Comme il est difficile de capter et de visualiser la structure de tableaux composés de nombreux chiffres, plusieurs méthodes (l'arbre de longueur minimum, l'analyse factorielle des correspondances, etc.), fondées sur des techniques d'analyse multidimensionnelle, sont utilisées pour construire des cartes qui rendent possibles diverses interprétations des données bibliométriques avec des finalités différentes.

Usages

Ces cartes de réseaux relationnels permettent de visualiser la structure de la recherche dans les différents champs et sous-champs de la science et d'observer, avec plus de clarté que dans les tableaux statistiques, l'ensemble des relations nouées entre les pays et/ou des domaines scientifiques. Grâce à ces techniques, il est possible de situer la position relative de pays de taille différente dans la collaboration scientifique mondiale.

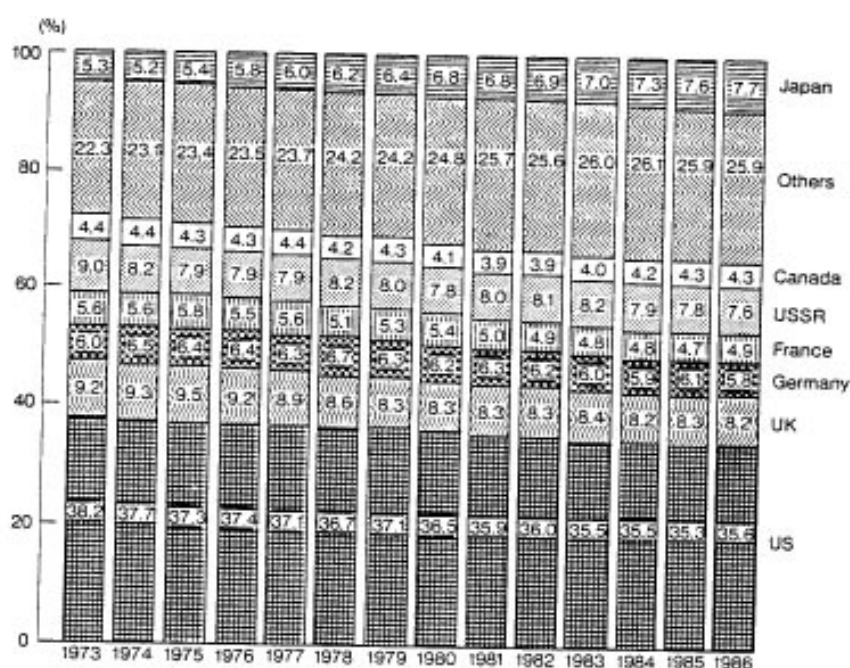
Limites

On ne peut pas représenter en deux dimensions toutes les données contenues dans un système multidimensionnel sans perdre des informations. Il est donc nécessaire de conjuguer diverses techniques pour atténuer ces pertes.

LISTE DES EXEMPLES

Exemple n° 1.

Evolution de la production scientifique mondiale 1973-86 par principaux pays (%)

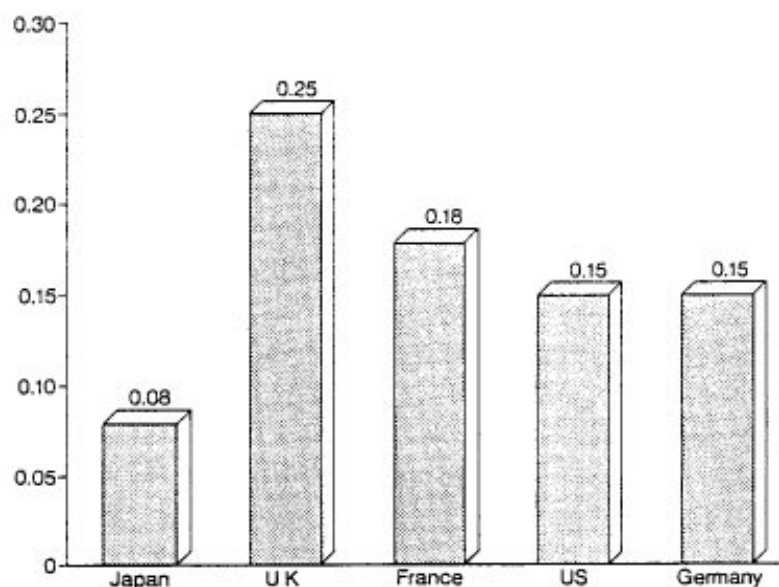


Source : Science and Technology Agency (Japon) (1991), p.13.

La ventilation de la littérature scientifique et technique mondiale par principaux pays “producteurs” fait partie des indicateurs bibliométriques “classiques”, comme en témoignent, par exemple, les différents rapports de la National Science Foundation (NSF) des États-Unis ou bien les séries de *Scientometrics* qui présentent, à des intervalles réguliers, ces informations ventilées par pays et par principaux domaines scientifiques.

Cette figure indique qu’en 1986 les publications des États-Unis représentent 35.6 pour cent de l’ensemble des publications mondiales dans les “sciences dures” et les “sciences de la vie” (c’est-à-dire à l’exception des “sciences molles”, comme les sciences sociales et humaines). Parmi les autres grands pays, le Royaume-Uni, le Japon et l’URSS comptent chacun pour environ de 8 pour cent du total. A l’exception du Japon et du groupe des “autres pays”, qui ont vu leur part croître au cours de la période, on observe une diminution relative pour tous les pays.

Exemple n° 2.
Le nombre de publications par chercheur dans cinq pays, 1986

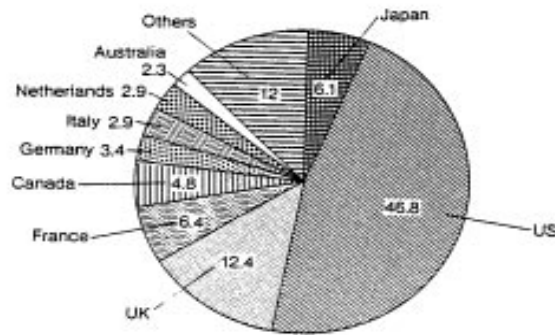


Source : Science and Technology Agency (Japon) (1991), p.13.

Ce graphique suggère que la “productivité”, mesurée par le nombre de publications par chercheur, est considérablement supérieure au Royaume-Uni et en France qu’au Japon, le ratio japonais n’étant qu’environ un tiers du chiffre britannique (0.08 au Japon contre 0.25 au Royaume-Uni). Pour relativiser ces informations, il convient de tenir compte d’un certain nombre de facteurs ayant un impact sur la comparabilité des données. Il faut noter que les statistiques de l’OCDE sur les chercheurs, utilisées pour le calcul de ces ratios, sont en effet surestimées pour le Japon (données plus proches du nombre total de personnes employées, alors qu’elles sont exprimées en équivalence plein-temps pour les autres pays). De plus, au Japon, une majorité des chercheurs travaillent dans l’industrie privée, tandis que la part des chercheurs universitaires est plus élevée dans la plupart des autres pays (les pratiques de publication étant assez différentes entre le secteur industriel et celui de l’enseignement supérieur). Finalement, comme l’analyse est fondée sur le nombre répertorié de publications en anglais, les références japonaises sont sous-représentées.

Exemple n° 3.

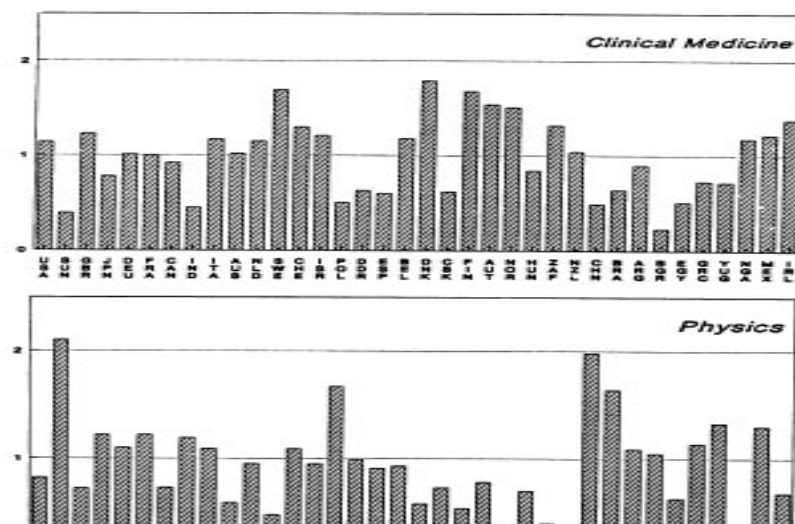
Ventilation par pays-auteurs des publications sur le thème “génomome humain”, 1991 (%)



Source : Science and Technology Agency (Japon) (1991), p.14.

Dans le domaine des sciences de la vie, les données sur le génome humain sont considérées comme une clé pour comprendre les fonctions organiques. Cette figure présente la distribution des pays-auteurs des publications sur ce sujet en 1991 (ventilation légèrement différente de celle de l'exemple n° 1). En comparant les contributions nationales, on trouve que les États-Unis et le Royaume-Uni sont les principaux acteurs, avec des parts d'ailleurs bien supérieures à celles de leur contribution à la production scientifique générale (voir l'exemple 1).

Exemple n° 4.
Spécialisation par discipline : le poids de la médecine clinique et de la physique dans les publications scientifiques nationales, 1981-86



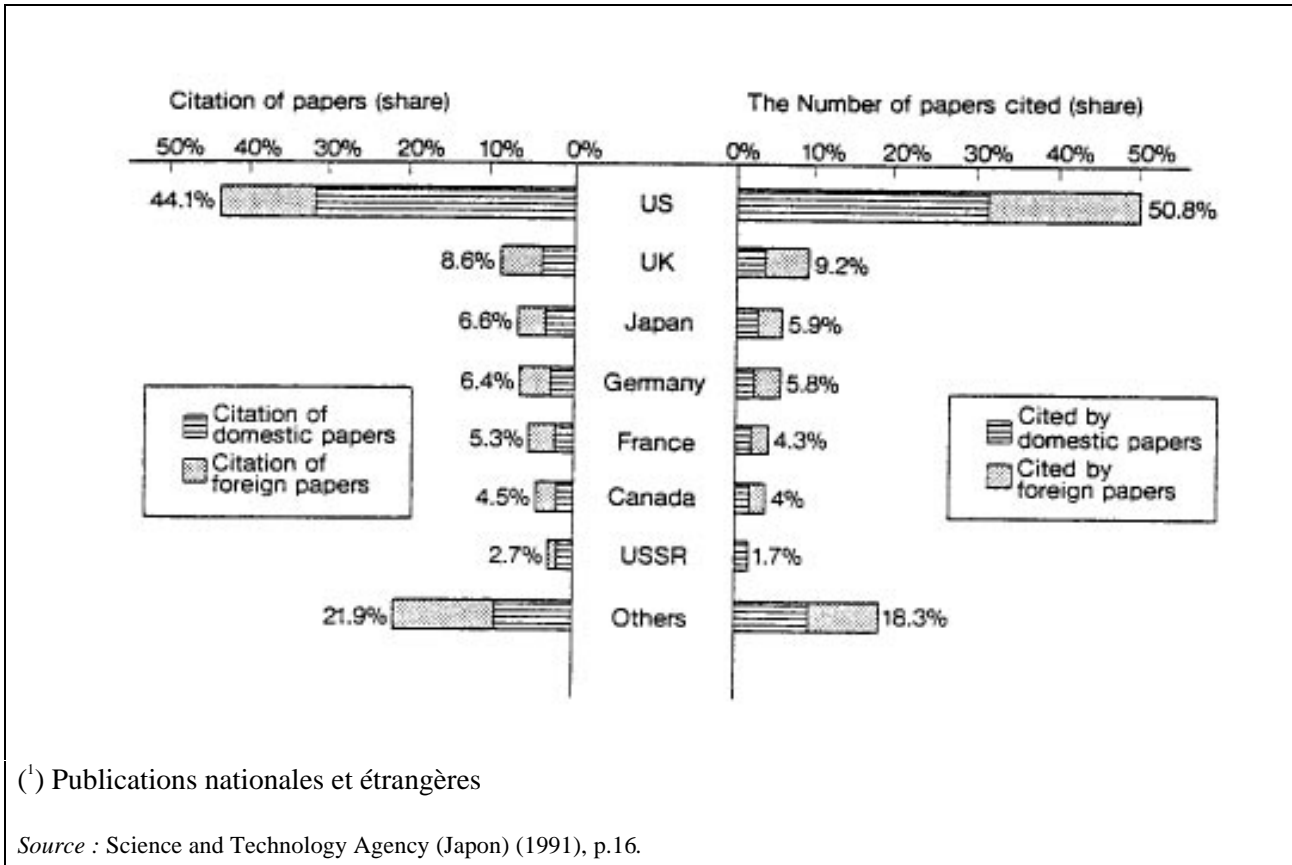
Source : Miquel et Okubo (1994), pp.271-297.
 Voir l'annexe pour les abréviations

La production nationale (mesurée par le nombre de publications scientifiques) dans les deux domaines de la médecine clinique et de la physique, vue par rapport à la proportion de ces mêmes domaines dans l'ensemble des publications enregistrées ("référence monde"), est présentée pour les 36 premiers pays "producteurs".

Le niveau 1 représente la "référence monde" qui est atteinte (ou ne l'est pas) par les pays ; cela illustre, en quelque sorte, leur "degré de spécialisation". Les pays s'approchant du niveau 2 (où le taux est le double de la "référence monde") ont donc des "spécialisations" très prononcées.

Ainsi, il y a d'importants contrastes, aussi bien entre les pays qu'entre les divers domaines scientifiques (domaines "forts" et "faibles"), par rapport au profil mondial des publications scientifiques. Ceci est particulièrement visible en médecine clinique et en physique. En médecine clinique, presque la moitié des pays ont une part proportionnelle plus élevée (typiquement, les pays scandinaves ou les pays culturellement liés au Royaume-Uni) que la "référence monde". L'histogramme pour la physique est, en quelque sorte, l'inverse de celui de la médecine clinique.

Exemple n° 5.
**Pourcentage des publications ⁽¹⁾ citant et étant citées par d'autres publications ⁽¹⁾,
 par principaux pays (moyennes 1984-86)**



Les citations de publications nationales et étrangères (référence monde) sont présentées pour un certain nombre de pays. De l'avis général, ce sont les publications étrangères qui ont le plus de signification (impact).

La partie gauche du graphique indique, par exemple, qu'environ 44 pour cent de l'ensemble des citations dans le monde sont faites par des scientifiques américains ; ceux-ci citent d'abord, pour environ les deux tiers, d'autres travaux américains, tandis que les chercheurs des autres pays citent relativement plus souvent des publications étrangères.

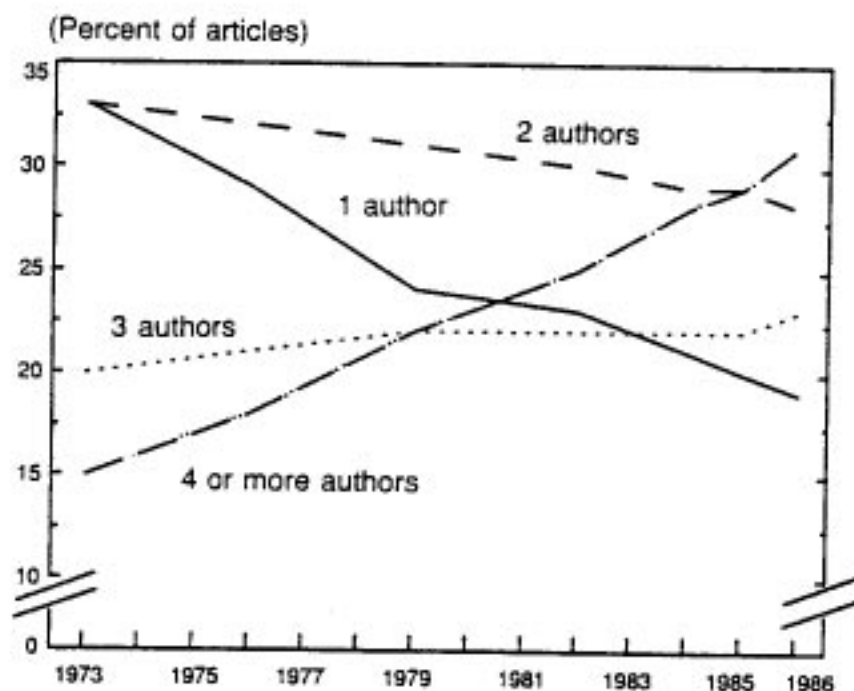
La partie droite du graphique indique que ce sont les publications des chercheurs américains qui sont les plus citées dans le monde (environ 51 pour cent) et, comme on vient de le dire, pour une très grande part par d'autres membres de la communauté scientifique américaine. Les scientifiques britanniques comptent pour quelque 9 pour cent des publications citées.

Exemple n° 6.
Citations parues dans des publications en 1984, ventilées par principaux domaines scientifiques et par l'orientation de R-D en ingénierie et technologie

Catégorie de la recherche					
	Total	Technologie appliquée	Ingénierie et science technologiques	Recherche appliquée et recherche fondamentale ciblée	Recherche fondamentale
Nombre de citations					
<i>Domaine cité</i>					
Ensemble des domaines	88 504	15 835	43 527	18 468	10 674
Ingénierie	59 483	15 093	40 599	3 660	139
Physique	14 501	441	647	8 787	4 626
Chimie	7 605	62	452	4 443	2 648
Autres	6 915	239	1 829	1 578	3 261
Pourcentage des citations					
Ensemble des domaines	100.0	17.9	49.2	20.9	12.1
Ingénierie	67.2	25.4	68.3	6.2	0.2
Physique	16.4	3.0	4.5	60.6	31.9
Chimie	8.6	0.8	5.9	58.4	34.8
Autres	7.8	3.5	26.4	22.8	47.2
Pourcentage des publications mondiales dans le domaine					
Domaine					
Ingénierie	100.0	41.5	50.6	7.9	0.0
Physique	100.0	1.0	2.4	32.5	64.1
Chimie	100.0	0.7	2.8	27.8	68.7
<p>Les publications reçoivent la catégorie des revues où elle paraissent ; le contenu de ces revues va de "recherche plus appliquée" à "recherche plus fondamentale".</p> <p>A l'exception de la colonne "Total" où les pourcentages se lisent verticalement, les colonnes indiquent les pourcentages horizontalement, c'est-à-dire elles indiquent le pourcentage des études dans chaque catégorie de recherche, par domaine de recherche.</p> <p>La distributions de l'ensemble mondial des publications a été déterminé pour les années 1973-79. Les données pour des années plus récentes suggèrent une distribution similaire.</p> <p>On a compté les citations de 1984 et des années précédentes (jusqu'en 1973). Les données représentent les articles, notes et comptes rendus dans plus de 2 100 des revues les plus influentes contenues dans <i>Science Citation Index Corporate Tapes</i> of the Institute for Scientific Information.</p> <p>Source : National Science Board (1989), p.346.</p>					

Les citations sont ici utilisées comme indicateur d'influence sur la recherche fondamentale dans le domaine de l'ingénierie et de la technologie ("sciences dures"). Le tableau présente (en chiffres absolus et en pourcentages) les citations mondiales cumulées depuis 1973. Les publications en question ont été classées dans des catégories selon leur orientation, allant de "plutôt recherche appliquée" à "plutôt recherche scientifique fondamentale" (colonnes) et par principaux sous-groupes d'ingénierie et de technologie (lignes). Il apparaît qu'en ingénierie et technologie, les publications citent de préférence leur propre sous-domaine scientifique. En 1984, environ 68 pour cent des citations en ingénierie étaient attribuées à d'autres publications du même champ, tandis qu'on observe des taux beaucoup plus bas (moins de 6 pour cent) pour la physique et la chimie. Les revues orientées plutôt vers la recherche fondamentale en ingénierie et technologie ne citent presque pas de publications du sous-groupe ingénierie (moins de 1 pour cent du total) mais beaucoup plus celles en physique et en chimie (environ un tiers du total chacun) et surtout (presque pour moitié, 47 pour cent) celles d'autres domaines.

Exemple n° 7.
Publications dans la science et l'ingénierie
ayant un, deux, trois, quatre ou plus d'auteurs 1973-86 (%)



Source : National Science Board (1989), p.120.

De 1973 à 1986, la moyenne du nombre d'auteurs par publication est passée, pour le monde entier, de 2.3 à 3.0. Simultanément, le pourcentage des publications ayant un seul auteur a décliné, de 33 pour cent du total en 1973 à 19 pour cent en 1986. Durant la même période, les publications co-signées de quatre auteurs (ou plus) sont passées de 15 à 31 pour cent du total. Cette tendance vers une augmentation du nombre de publications co-signées ne peut être qu'un signe d'une plus grande coopération entre chercheurs et groupes de recherche, c'est-à-dire du développement d'équipes de recherche plus importantes ou de recherches liées à de grands programmes nationaux et internationaux (big science). Un examen plus approfondi des données (voir les statistiques des publications de la NSF) révèle toutefois d'assez grandes variations entre les différentes disciplines. Ainsi, c'est dans les domaines de la médecine clinique et de la recherche biomédicale que l'on trouve les taux de coopération les plus élevés (avec, en 1986, 43 et 34 pour cent, respectivement, des publications ayant quatre coauteurs ou plus et seulement 14 pour cent un seul auteur). En revanche, dans les mathématiques, presque les deux tiers (62 pour cent) des publications sont l'oeuvre d'un seul auteur, 30 pour cent de deux auteurs, 7 pour cent de trois auteurs. Un pour cent seulement du total des publications en mathématiques résulte d'une coopération entre quatre auteurs ou plus. Le graphique présente les tendances de coopération scientifique dans le monde depuis le début des années 70, tous domaines scientifiques confondus.

Exemple n° 8.
Publications co-signées université-industrie aux États-Unis, par rapport à la totalité des articles signés des chercheurs industriels, par domaines scientifiques 1981-91 (%)

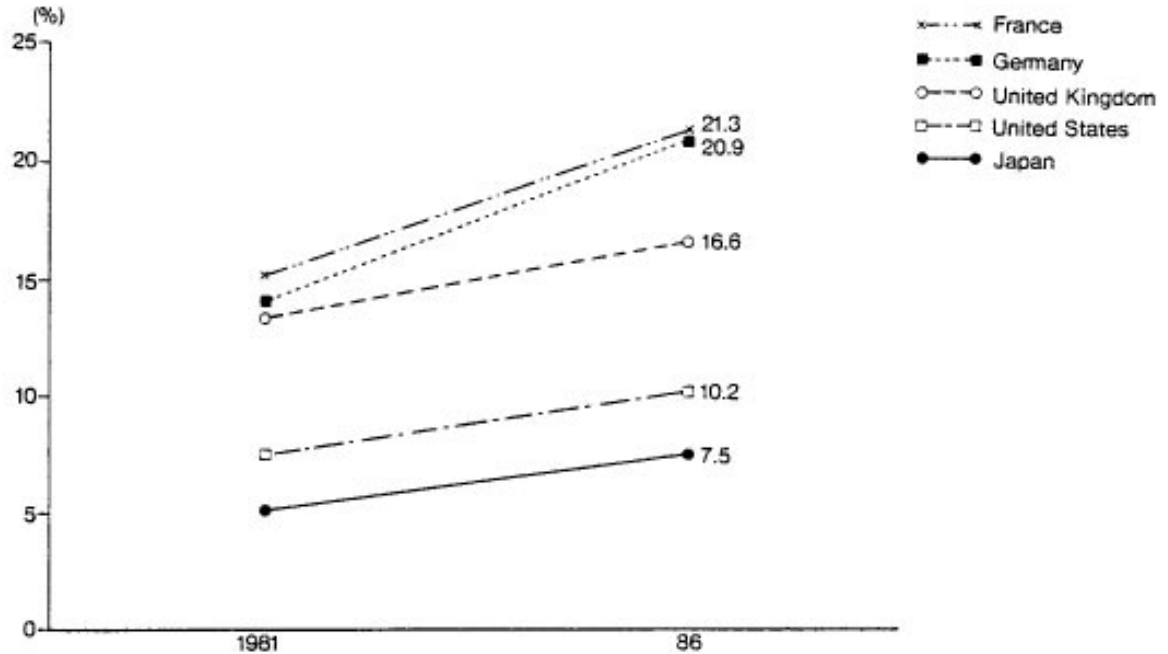
	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
Ensemble des domaines	22	24	23	25	27	28	30	31	32	33	35
Médecine clinique	30	34	33	35	40	37	42	41	42	44	45
Recherche biomédicale	35	37	35	35	39	38	40	41	39	39	40
Biologie	39	46	42	37	44	44	41	47	48	43	45
Chimie	13	17	15	16	16	18	20	20	22	22	24
Physique	20	21	23	25	23	23	25	26	28	29	31
Sciences de la Terre et de l'espace	34	35	33	36	33	36	34	41	38	40	37
Ingénierie et technologie	16	17	16	17	18	20	23	24	23	26	26
Mathématiques	43	35	42	42	43	40	42	41	51	52	49

Source : National Science Board (1989), p.429.

Ce tableau présente l'évolution de la coopération université - industrie aux États-Unis au cours de la décennie 1981-91 mesurée en termes de publications co-signées, par grands domaines scientifiques.

Cette coopération est en nette croissance. En 1991, 35 pour cent de toutes les publications signées par des industriels avaient des co-signataires universitaires (contre 22 pour cent en 1981). Cette tendance affecte toutes les disciplines, à des degrés divers. Ce sont les mathématiques qui ont le pourcentage de coopération le plus élevé (presque la moitié de toutes les publications) ainsi que les sciences de la vie (médecine clinique, biologie et recherche biomédicale) avec des taux autour de 40 à 45 pour cent. Les publications dans les domaines de la chimie et de l'ingénierie-technologie sont celles où les pourcentages sont les plus faibles (respectivement 24 pour cent et 26 pour cent).

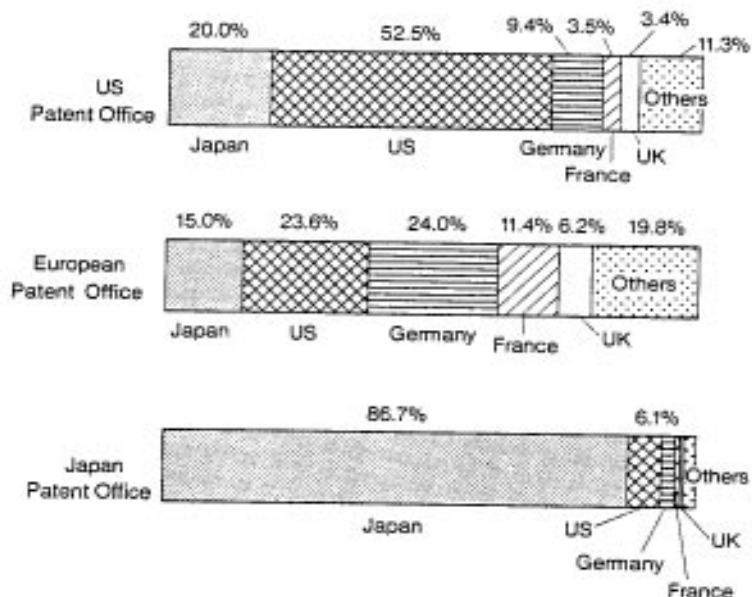
Exemple n° 9.
Tendances des co-signatures internationales de quelques pays, 1981-86 (%)



Source : Science and Technology Agency (Japon) (1991), p.17.

Les données montrent une croissance de l'internationalisation de la recherche. Entre 1981 et 1986, les grands pays ont accru le pourcentage des publications nationales co-signées avec des étrangers. En 1986, environ 10 pour cent de toutes les publications scientifiques aux États-Unis étaient co-signées avec des auteurs ayant une affiliation étrangère. Les pays de l'Europe de l'Ouest se situaient à la même époque entre 15 et 20 pour cent. C'est le Japon qui avait le taux le plus faible d'internationalisation, ainsi mesurée.

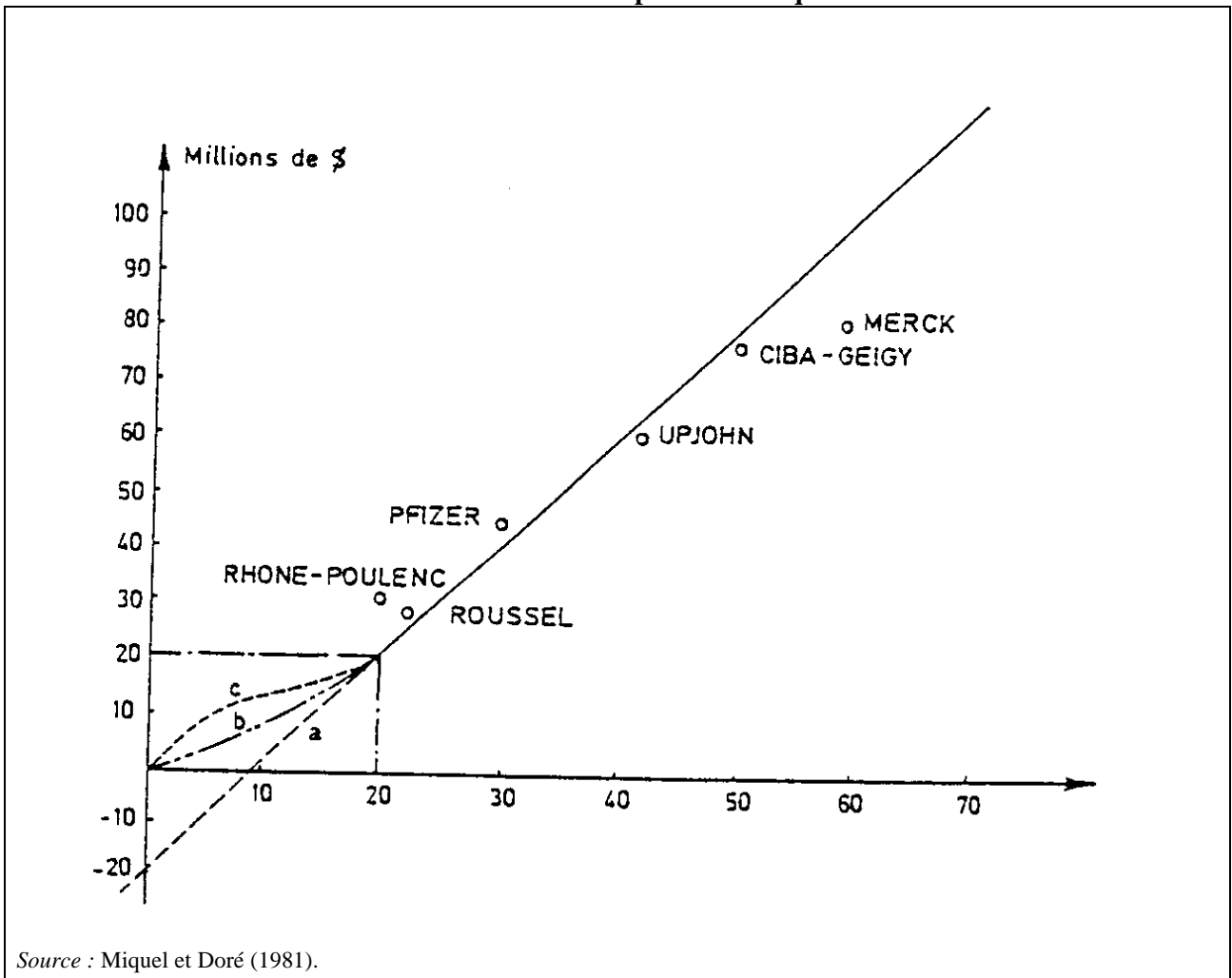
Exemple n° 10.
Ventilation des brevets déposés auprès des bureaux des brevets des États-Unis, de l'Europe et du Japon, par principaux pays (%)



Source : Science and Technology Agency (Japon) (1991), p.13.

Aux États-Unis, les dépôts de brevets résidents représentent environ la moitié (52.5 pour cent) de l'ensemble des demandes vers la fin des années 80. Ce taux est bien plus élevé au Japon, où les demandes des résidents comptent pour une très large majorité (près de 87 pour cent) et celles des non-résidents pour environ 13 pour cent du total. Un cinquième des demandes enregistrées auprès du Bureau américain des brevets (US Patent Office) et 15 pour cent de celles enregistrées auprès du bureau européen proviennent du Japon. Les États-Unis semblent bien plus enclins à breveter en Europe qu'au Japon. Le taux de dépôts de brevets au Japon des principaux pays européens est très inférieur à celui de leurs dépôts aux États-Unis.

Exemple n° 11.
“Productivité” de la recherche pharmaceutique industrielle



Ce graphique présente, pour quelques-uns des principaux groupes pharmaceutiques multinationaux, la “productivité” de leurs activités scientifiques, exprimée par le rapport entre chiffres d’affaires (axe vertical) et nombre de brevets accordés (axe horizontale) au cours des années 70; les brevets sont ici considérés comme une mesure approximative des “résultats” de la recherche.

La droite de régression de la figure met en évidence le fait qu’à partir d’un certain seuil, une corrélation relativement satisfaisante ($r = 0.814$) existe entre les deux variables, ce qui montre que la politique des brevets suit une “loi” étroitement liée au budget de recherche. Les différentes entreprises suivent cette “loi” de façon similaire, car elles ne peuvent pas se permettre de trop grands écarts dans la gestion de leurs recherches, dans un marché à forte compétition où la production scientifique conditionne non seulement les profits à venir, mais la survie même.

Exemple n° 12.

Citations des brevets américains par des brevets postérieurs, classés par pays inventeurs et secteurs propriétaires des brevets cités

a. Pays de l'inventeur										
Année de brevet cité	États-Unis	Japon	Pays-Bas	Royaume-Uni	Allemagne de l'Ouest	Canada	France	Suisse	Italie	Suede
Citations par brevet citable										
1975	3.91	4.03	3.63	3.63	3.33	3.33	3.09	2.98	2.99	3.24
1980	3.09	3.35	2.90	2.70	2.51	2.50	2.41	2.37	2.34	2.28
1985	1.37	1.84	1.15	1.21	1.15	1.11	1.07	1.10	1.04	1.02
b. Secteur du propriétaire pour les inventeurs des États-Unis										
	Tous inventeurs	Sociétés	Gouvernement	Individus	Étrangers					
1975	3.91	4.15	3.00	3.34	4.13					
1980	3.09	3.37	2.41	2.46	3.42					
1985	1.37	1.49	1.07	1.02	1.44					
Les chiffres, surtout ceux pour les dernières années, augmenteront à mesure que les brevets continueront à être cités.										
<i>Source : National Science Board (1989), p.140..</i>										

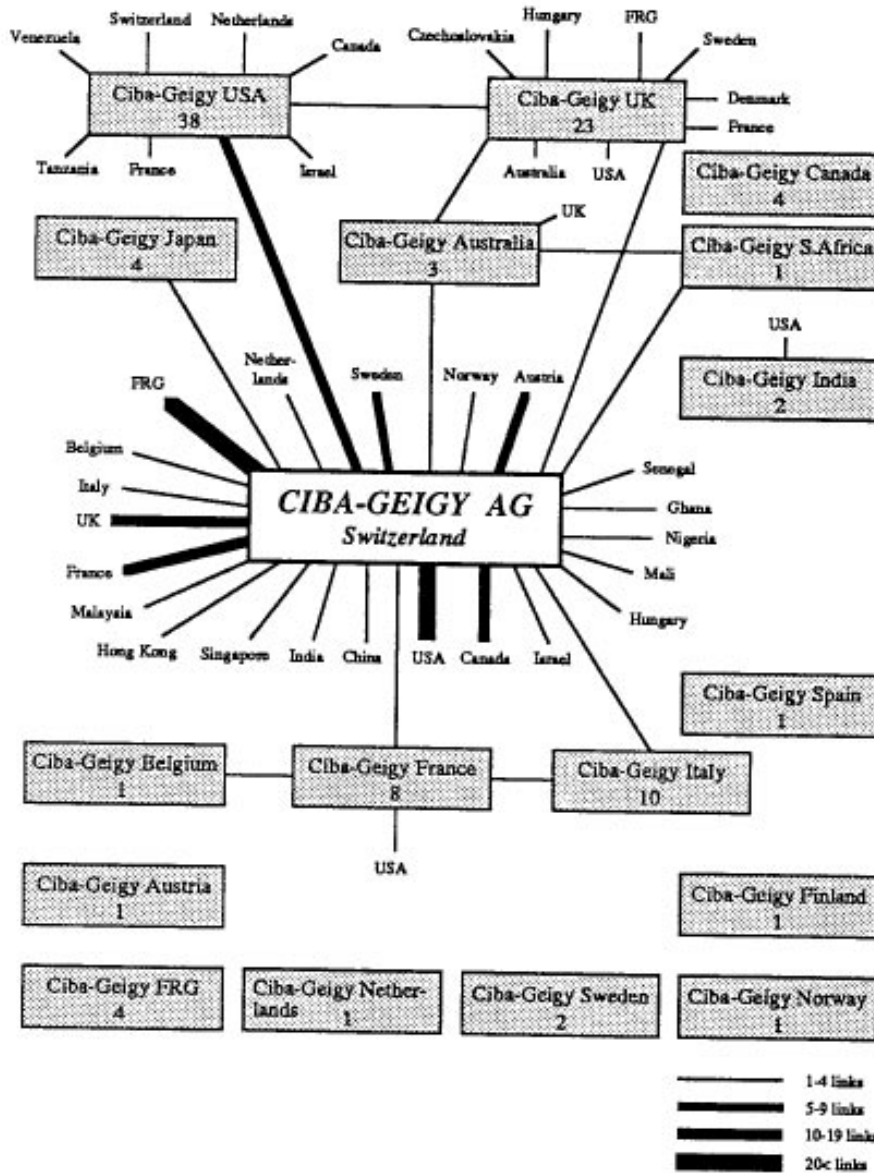
Les données concernent les citations que des brevets déposés aux États-Unis font, au cours d'une année donnée, à d'autres brevets. Il est possible de voir les pays (et également, dans les tableaux détaillés de la NSF, les produits ou branches industrielles) dont les brevets reçoivent le plus ou le moins de citations.

Il convient toutefois de noter que la "structure" de ces citations peut varier considérablement, non seulement entre secteurs industriels mais aussi au fil du temps, du fait des "décalages temporels". Autrement dit, plus un brevet est "ancien", plus il aura eu le temps de faire l'objet de citations, et plus il est "jeune", moins son importance aura attiré l'attention d'autres déposants de brevets.

Parmi les dix pays étudiés, le Japon est celui dont les brevets sont le plus cités au cours de la période étudiée, suivis par les brevets américains, ceux des Pays-Bas, du Royaume-Uni et de l'Allemagne. Ceci semble en accord avec le nombre absolu des brevets déposés par ces mêmes pays aux États-Unis.

En ce qui concerne la ventilation par secteur du propriétaire des brevets cités, ce sont les entreprises américaines qui possèdent les brevets les plus cités. Ceux qui appartiennent au gouvernement des États-Unis, ou à des Américains en tant qu'individus, sont les moins cités. En 1975 et en 1980, les compagnies américaines ont eu autant de citations de leurs brevets que les propriétaires de brevets japonais. Presque tous les brevets japonais déposés aux États-Unis appartiennent à des entreprises.

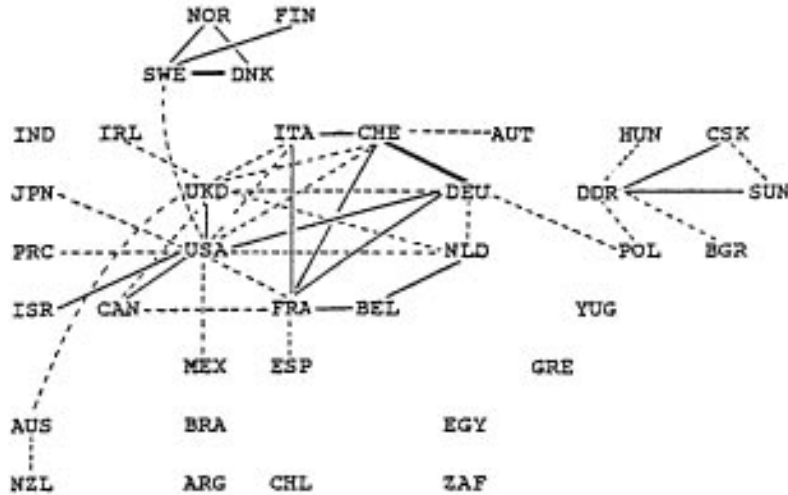
Exemple n° 13.
Le réseau international de la Ciba-Geigy Corporation, 1990



Source : Okubo (1993).

Les rectangles gris indiquent les filiales qui ont participé à l'activité scientifique et les chiffres le nombre de publications. Ciba-Geigy, une société pharmaceutique, montre un degré d'internationalisation élevé, puisque quelque 28 pour cent des publications scientifiques du groupe se font en co-signatures internationales et 32 pays sont en relation directe avec les centres d'excellence de Ciba-Geigy, à travers la maison-mère en Suisse ou l'une de ses filiales. Le degré d'internationalisation des recherches, par rapport à la production scientifique totale du groupe, est plus élevé pour la maison-mère (dont 36 pour cent des publications sont co-signées). Cependant, les filiales aux États-Unis et au Royaume-Uni participent également à cette internationalisation en créant leurs propres liens internationaux avec des universités et hôpitaux étrangers.

Exemple n° 14.
Carte des liens de coopération internationale dans les sciences, 1981-85



=== indique 2.5 % $< r_{ik}$
 _____ indique 1.5 % $< r_{ik} < 2.5$ %
 - - - - - indique 1.0% $< r_{ik} < 1.5$ %
 La puissance relative r_{ik} est donnée par $r_{ik} = n_{ik}/(n_i \cdot n_k)^{1/2}$ quand n_i et n_k = le nombre total des articles publiés par les pays I et k respectivement.

Source : Schubert et Braun (1990).

Ce graphique trace les principaux réseaux internationaux de coopération scientifique pour les 36 pays les plus "productifs" parmi les 163 pays du monde ayant produit, au cours de la période, au moins une publication en coopération internationale. Ce graphique a été établi à partir du tableau n°1 de l'article, cité comme source, qui indique, pour tous ces pays, le nombre total de leurs publications ainsi que le nombre de celles produites en coopération ; cela permet le calcul de ratios, exprimés en pourcentages, qui indiquent le degré d'internationalisation. Les auteurs observent, pour ces 36 pays, quatre grands agrégats (*clusters*) de taille différente ; le plus grand couvre les États-Unis et le Canada ainsi que la plupart des pays de l'Europe de l'Ouest, le second couvre les pays nordiques, le troisième les principaux pays de l'Europe de l'Est et le quatrième, très petit, l'Australie et la Nouvelle-Zélande. Les liens du groupe des pays de l'Est avec le monde scientifique occidental semble passer par la Pologne et ceux du groupe nordique, avec les États-Unis notamment, par la Suède. L'Australie et la Nouvelle-Zélande ont des liens prononcés avec le Royaume-Uni, et Israël avec les États-Unis. Bon nombre de ces configurations semblent refléter des liens historiques et/ou politiques.

Exemple n° 15.
Position du Japon dans l'activité scientifique internationale
de l'Allemagne, de la France et du Royaume-Uni
en ingénierie (ENT) et sciences de la Terre et de l'espace (EAS), 1981-86

<i>Allemagne/Germany</i>		<i>France</i>		<i>Royaume-Uni/United Kingdom</i>	
ENT	EAS	ENT	EAS	ENT	EAS
ÉTATS-UNIS	ÉTATS-UNIS	ÉTATS-UNIS	ÉTATS-UNIS	ÉTATS-UNIS	ÉTATS-UNIS
SUISSE	ROYAUME-UNI	ALLEMAGNE	ALLEMAGNE	CANADA	AUSTRALIE
FRANCE	FRANCE	CANADA	ROYAUME-UNI	ALLEMAGNE	ALLEMAGNE
ROYAUME-UNI	PAYS-BAS	ROYAUME-UNI	CANADA	AUSTRALIE	CANADA
JAPON	ITALIE	ITALIE	ITALIE	FRANCE	FRANCE
JAPON	CANADA	POLOGNE	PAYS-BAS	CHINE	PAYS-BAS
	SUEDE	BELGIQUE	URSS	EGYPTE	ITALIE
	JAPON	SUISSE	SUISSE	PAYS-BAS	AFRIQUE DU SUD
		EGYPTE	ESPAGNE	JAPON	NORVEGE
		ESPAGNE	BELGIQUE		INDE
		JAPON	SUÈDE		SUÈDE
			JAPON		GRÈCE
					ESPAGNE
					SUISSE
					JAPON
<i>% du Japon</i>					
4.5	2.6	2.0	1.8	2.5	1.2

Source : Okubo (1994).

Ce tableau situe le Japon dans l'ensemble des activités internationales, en termes du nombre des liens scientifiques établis, des trois principaux pays européens dans deux grands domaines scientifiques où la puissance japonaise est contrastée. L'Allemagne, la France et le Royaume-Uni ont chacun leur spécificité en matière d'activité internationale ; d'où des stratégies contrastées, soit explicites soit implicites, dans l'édification ou l'acquisition des connaissances. Il en ressort qu'en Allemagne, l'importance du Japon est nettement plus marquée en ingénierie que dans les sciences de la Terre et de l'espace. Le Japon est mieux placé dans le domaine où il est fort. Il y a donc une certaine "rationalité" dans le mode de coopération entre l'Allemagne et le Japon.

On retrouve, dans une certaine mesure, cette corrélation dans la coopération du Royaume-Uni, le Japon est mieux placé, ici aussi, en ingénierie qu'en Terre-espace. Toutefois, en ingénierie, le Royaume-Uni collabore beaucoup plus, en volume, avec le Canada, l'Australie, la Chine et l'Égypte qu'avec le Japon. L'expérience historique semble jouer ici un rôle plus important pour l'établissement des liens scientifiques. La France n'adopte pas la même logique "rationnelle" que l'Allemagne et le Japon. En ce qui concerne le Japon, il n'y a aucune différence significative entre ingénierie et terre-espace ; la France travaille d'abord avec ses voisins.

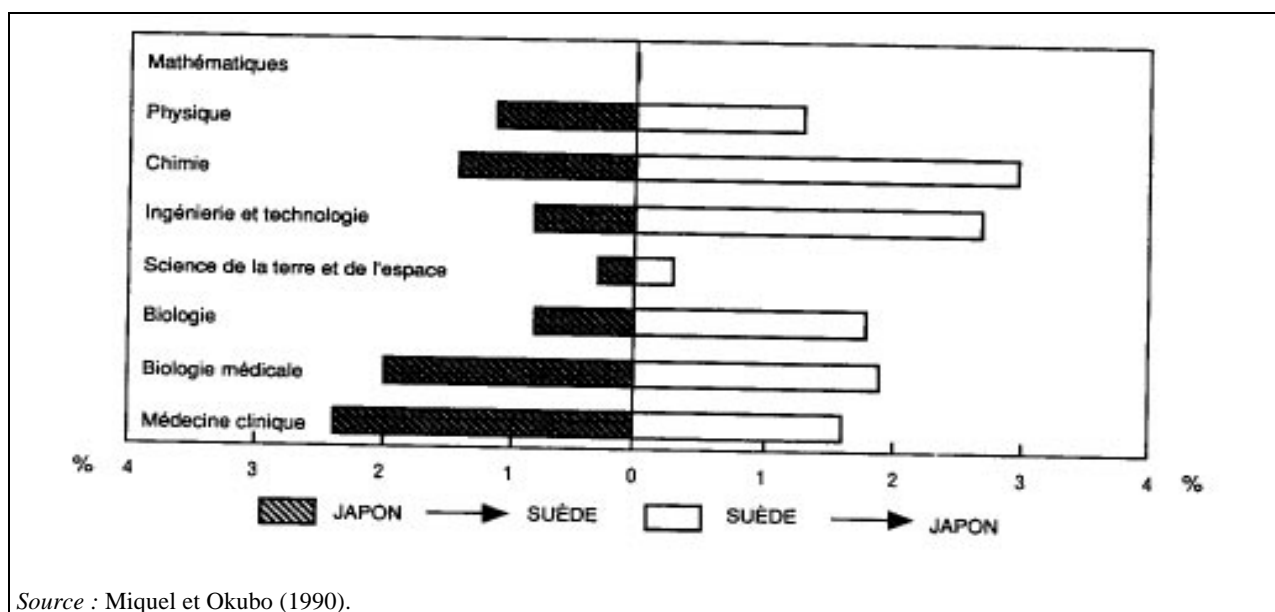
Exemple n° 16.
Profils comparés – Strasbourg et France, 1986
Liens de coopération scientifique entre Strasbourg et 97 pays partenaires

Pays	MAT	PHY	CHM	ENT	EAS	BIO	BIM	CLI	Total
États-Unis		16	12	3	3	3	34	4	75
Allemagne		26	2	1	2	2	23	7	63
Royaume-Unis	1	9	8		1	4	17	2	42
Italie	1	11	10			2	8	6	38
Canada		1	3			1	6		11
Suisse		17			2	2	14	5	40
Belgique	1	7	3			1	4		16
Espagne	1	10	10			1	4		26
Pays-Bas		3	2				5		10
Suède		6	1				5	2	14
...	
Autres					1		3	1	5
Nbre de liens	4	158	57	11	20	21	142	28	441
Profil de Strasbourg (a)	0.9	35.8	12.9	2.5	4.5	4.8	32.2	6.4	100%
Profil de la France (b)	2.7	30.7	11.8	3.2	8.3	4.3	18.0	21.0	100%
Profils/comparés (a/b)	0.3	1.2	1.1	0.8	0.5	1.1	1.8	0.3	
MAT (mathématiques), PHY (physique), CHM (chimie), ENT (ingénierie-technologie), EAS (terre-espace), BIO (biologie fondamentale et appliquée), BIM (biomédicale), CLI (médecine clinique)									
<i>Source : Okubo et al. (1992).</i>									

La ville de Strasbourg est un des principaux pôles scientifiques de la France. En comparant les liens de Strasbourg avec ceux de la France, on peut identifier les particularités des activités internationales de cette ville universitaire. Strasbourg a une part proportionnellement importante des liens de coopération internationale en physique : 158 liens (ou 35.8 pour cent de l'ensemble des liens de la ville). Dans le reste de la France la physique compte pour 30.7 pour cent de l'ensemble des liens. Si l'on compare la place de la physique dans les relations internationales de Strasbourg à la place qui lui est faite pour toute la France ("profil comparé"), on remarque que les chiffres sont légèrement supérieurs à Strasbourg (x1.2). La participation d'institutions strasbourgeoises à des projets du CERN est l'une des raisons de cette inclinaison, car le CERN contribue pour 23 pour cent aux articles strasbourgeois co-signés en physique. Ces activités de physique nucléaire touchent aussi l'ingénierie et la technologie, par le biais de méthodes et de construction d'appareillages spécifiques.

Une part exceptionnelle d'activités internationales relève de la recherche biomédicale (142 liens ou 32.2 pour cent du total) et représente un autre trait distinctif de la science strasbourgeoise ; ce pourcentage est supérieur à la moyenne du pays (18 pour cent). Le niveau des collaborations est dû, d'une part, à la présence de plusieurs laboratoires spécialisés (biologie moléculaire-ingénierie génétique, biologie moléculaire cellulaire, neurobiologie, etc.), d'autre part, à la présence de deux institutions industrielles qui participent à la coopération internationale biomédicale. En revanche, la proportion de la médecine clinique dans les liens de Strasbourg (6.4 pour cent) est particulièrement basse, aussi bien par rapport aux activités internationales globales de la ville que par rapport à la médecine clinique totale en France.

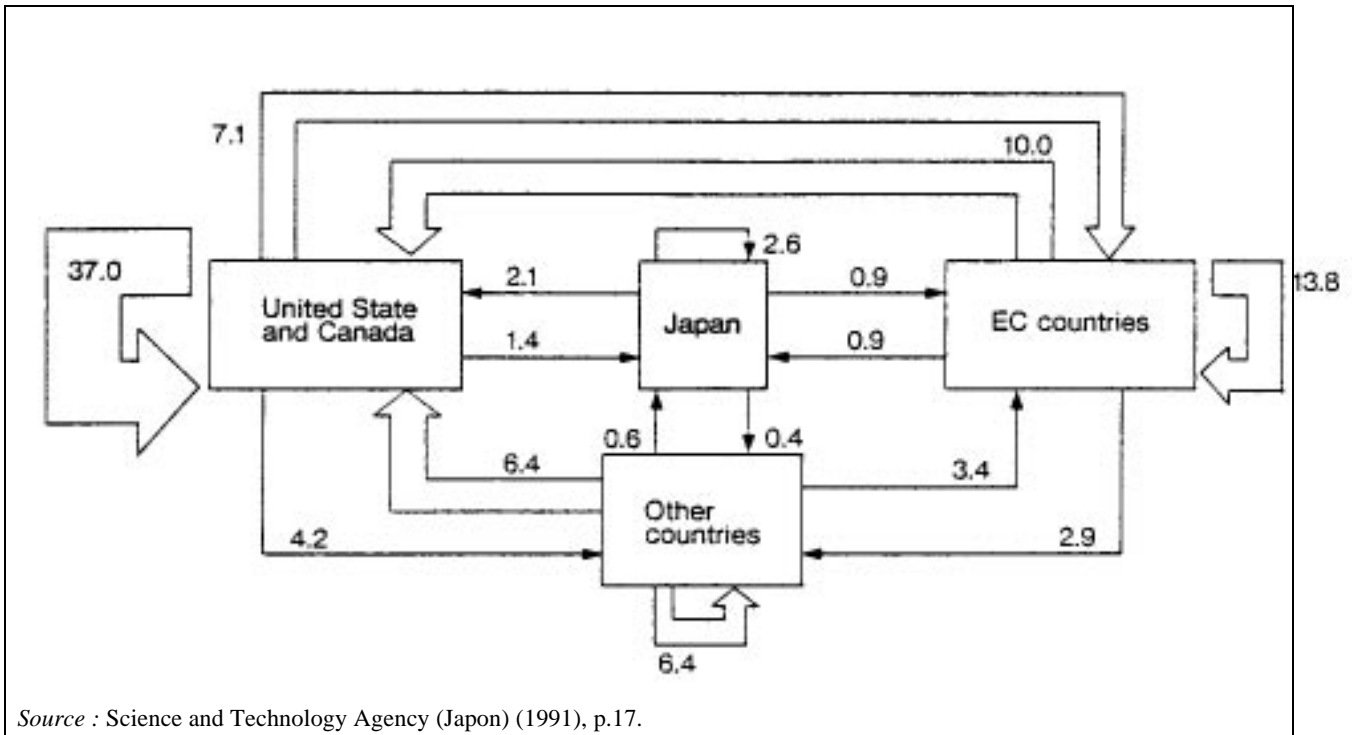
Exemple n° 17.
Indices d'affinité entre le Japon et la Suède 1981-86 (%)



Ce graphique présente l'utilisation de l'indice d'affinité pour illustrer la coopération scientifique entre le Japon et la Suède.

Les liens de collaboration semblent indiquer que chacun favorise les domaines scientifiques où l'autre est le plus fort. La coopération japonaise avec la Suède est particulièrement élevée en médecine clinique et en biologie médicale - les domaines de recherche prédominants en Suède - alors que la collaboration suédoise avec le Japon atteint son maximum en chimie et ingénierie et technologie, domaines où les Japonais sont relativement forts. Au contraire, il n'y a guère de collaboration en mathématiques ou dans les sciences de la Terre et de l'espace, domaines où ni l'un ni l'autre n'a des compétences exceptionnelles.

Exemple n° 18.
Répartition des citations dans le monde, 1984-88 (%)

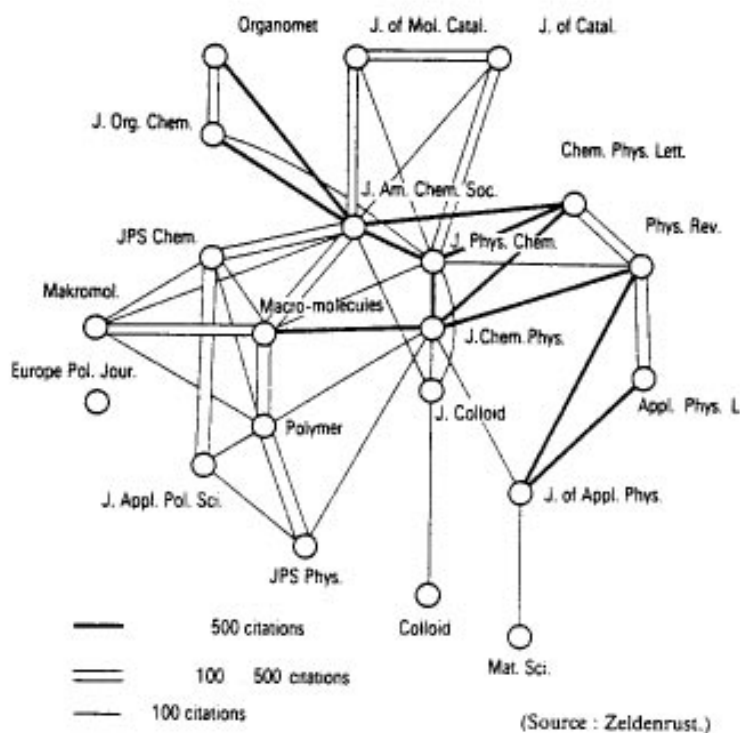


Ce graphique illustre la répartition du total mondial des citations (tous domaines scientifiques confondus) d'articles scientifiques pendant la période 1984-88, ventilée selon les principaux (groupes de) pays des citations. Le graphique essaie de répondre aux questions : "Qui cite qui, qui est cité par qui, et y a-t-il des (dés)équilibres entre différents pays dans ces citations ?"

Il en ressort, par exemple, que l'Amérique du Nord (les États-Unis et le Canada) "reçoivent" quelque 55.5 pour cent des citations mondiales (dont 37 pour cent se réfèrent à des publications de la région même), 10 pour cent proviennent des pays de la Communauté européenne, 2.1 pour cent du Japon et 6.4 pour cent d'autres pays. De même, l'Amérique du Nord bénéficie de près de la moitié des citations mondiales (49.7 pour cent) dont 7.1 pour cent (citations "intra-muros" à part) portent sur des articles des pays de l'UE, 4.2 pour cent sur ceux d'autres pays et 1.4 pour cent sur le Japon.

On notera que la banque de données de la base SCI permet l'identification des liens scientifiques bilatéraux, avec une ventilation plus détaillée, par sous-disciplines désagrégées.

Exemple n° 19.
Citations entre revues spécialisées du domaine des polymères



Source : Callon *et al.* (1993), pp.62.

Cette présentation graphique résume les liens, mesurés en termes de citations, entre les principaux spécialisés dans le domaine des polymères.

“L’épaisseur des traits donne une idée de l’intensité des liens, c’est-à-dire du nombre de citations échangées entre articles de ces revues. Plus les revues sont liées, plus les contenus sont complémentaires et plus l’agrégat (“cluster”) qu’elles forment trahit l’existence d’une discipline cohérente et intégrée. L’analyse dynamique de ces agrégats permet de suivre la déformation et les réorganisations de ces domaines de recherche. Il faut tenir compte du fait que des revues peuvent voir le contenu de leurs articles se modifier au cours du temps et que la stabilité des agrégats peut bien cacher des changements thématiques importantes.” (Callon *et al.*, 1993, p.63)

Exemple n° 20.
Relation à la science des brevets du monde entier déposés aux États-Unis

Activité	Intensité scientifique des brevets		Age articles cités (années)
	(En nombre)	(Moyenne = 100)	
Electronique	0.31	29	7.7
Pharmacie	3.19	298	7.8
Instruments	0.54	51	8.1
Machines	0.12	11	8.2
Transports*	0.01	1	8.3
Chimie	0.77	72	9.7
Autres	0.12	11	8.5
Moyenne arithmétique	1.07	100	-
Ensemble	0.40	-	-

*Transports terrestres et aériens

Source : Observatoire des Sciences et des Techniques (1992), pp.142-143.

Le tableau présente une ventilation de “l’intensité scientifique” des brevets, classée par l’activité industrielle et technologique de leurs déposants. Il s’agit de références aux publications scientifiques du SCI qui font l’objet d’une citation par un ou plusieurs brevets. “La pharmacie est l’activité dont la technologie est la plus proche de la science, suivie de la chimie, des instruments et de l’électronique. La technologie de l’activité transports, à l’inverse, ne fait pratiquement pas référence à des publications scientifiques. En termes de temps nécessaire au passage science-technologie, l’électronique et la pharmacie se distinguent par une rapidité particulière ; en revanche, la chimie s’appuie en moyenne sur de la science plus ancienne.” (OST, 1992, chapitre 3)

Exemple n°21.

Répartition par discipline scientifique des citations scientifiques dans les dépôts de brevets du monde entier aux États-Unis (moyennes, 1987-88)

Discipline scientifique	Citations par les brevets (%)
Médecine clinique	20.0
Recherche biomédicale	20.5
Biologie animale et végétale	2.5
Chimie	23.0
Physique	18.0
Géosciences-espace	0.0
Sciences pour l'ingénieur	15.5
Mathématiques	0.0

Source : Observatoire des Sciences et des Techniques (1992), pp.142-143.

Ce tableau (basé sur les mêmes statistiques que l'exemple précédent) indique la part des différentes disciplines dans les citations faites par les brevets à des publications scientifiques. "La répartition des disciplines scientifiques des publications citées par les brevets montre que la médecine clinique, la recherche biomédicale, la chimie, la physique et les sciences de l'ingénieur se répartissent presque également les citations. La biologie animale et végétale, les géo-sciences/espace et les mathématiques sont des disciplines auxquelles les brevets ne font pas ou pratiquement pas référence." (OST, 1992, chapitre 3)

Exemple n° 22.
Production et utilisation de science en rapport avec les brevets
(moyenne 1987-88) en poids mondial

	Science	Production Science à potentiel technologique	Science utilisée	Brevets	Citation science
France	4.9	4.8	3.3	3.7	2.9
Allemagne	6.0	6.6	3.8	10.3	5.2
Royaume-Uni	8.9	8.1	7.5	3.6	3.8

Source : Observatoire des Sciences et des Techniques (1992), pp.142-143.

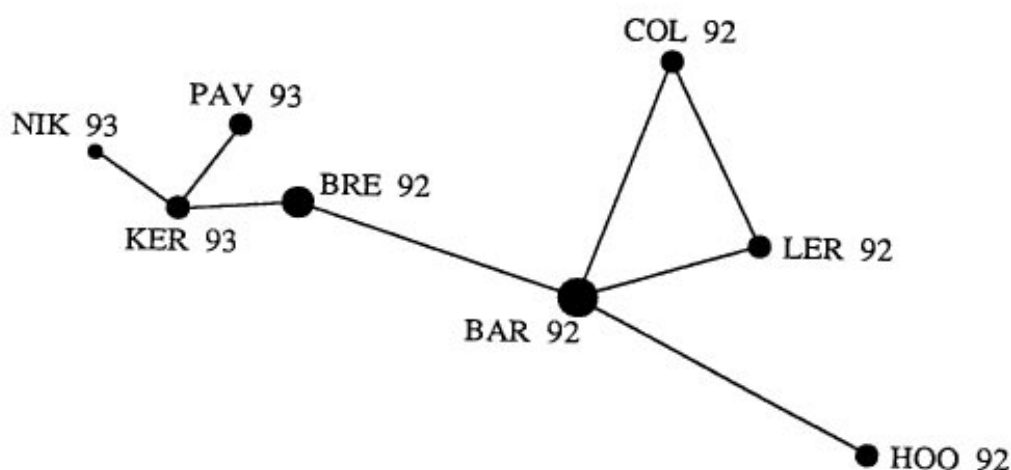
Toujours basé sur les mêmes statistiques, ce tableau montre pour la France, l'Allemagne et le Royaume-Uni les paramètres (indicateurs) suivants (tous définis en pourcentage du total mondial) :

- la production scientifique : la part mondiale des trois pays dans les publications scientifiques (toutes disciplines confondues) ;
- la production scientifique à potentiel technologique (= science susceptible d'être citée par les brevets) : la part mondiale des pays dans les publications scientifiques en moyenne des disciplines, pondérées par leur potentiel technologique ;
- la production scientifique "utilisée" par les brevets : la part mondiale des pays dans les publications scientifiques citées par l'ensemble des brevets déposés (tous pays confondus) ;
- la production de brevets : la part mondiale des pays dans l'ensemble des brevets déposés aux États-Unis ;
- l'utilisation de la science dans les brevets : part mondiale des pays dans les citations, faites par les brevets, de publications scientifiques (nombre de citations de publications scientifiques dans les brevets du pays, rapporté aux citations dans les brevets du monde entier).

"La production scientifique allemande à potentiel technologique n'est que peu inférieure à celle du Royaume-Uni (6.6 pour cent du total mondial contre 8.1 pour cent) mais elle est beaucoup moins 'utilisée' (citée par les brevets) que la science britannique (3.8 pour cent du total mondial contre 7.5 pour cent). La production scientifique française est presque aussi 'utilisée' pour la création technologique mondiale que la science allemande (3.3 pour cent contre 3.8 pour cent).

"C'est l'Allemagne, cependant, qui 'utilise' le plus de science dans l'absolu pour sa technologie (5.2 pour cent) ; ceci reflète le poids technologique important de l'Allemagne (10.3 pour cent), soit près de trois fois le Royaume-Uni et la France." (OST, 1992, chapitre 3).

Exemple n° 23.
Agrégats (“clusters”) de co-citations en 1992 et 1993
dans le domaine de la “chimie combinatoire”



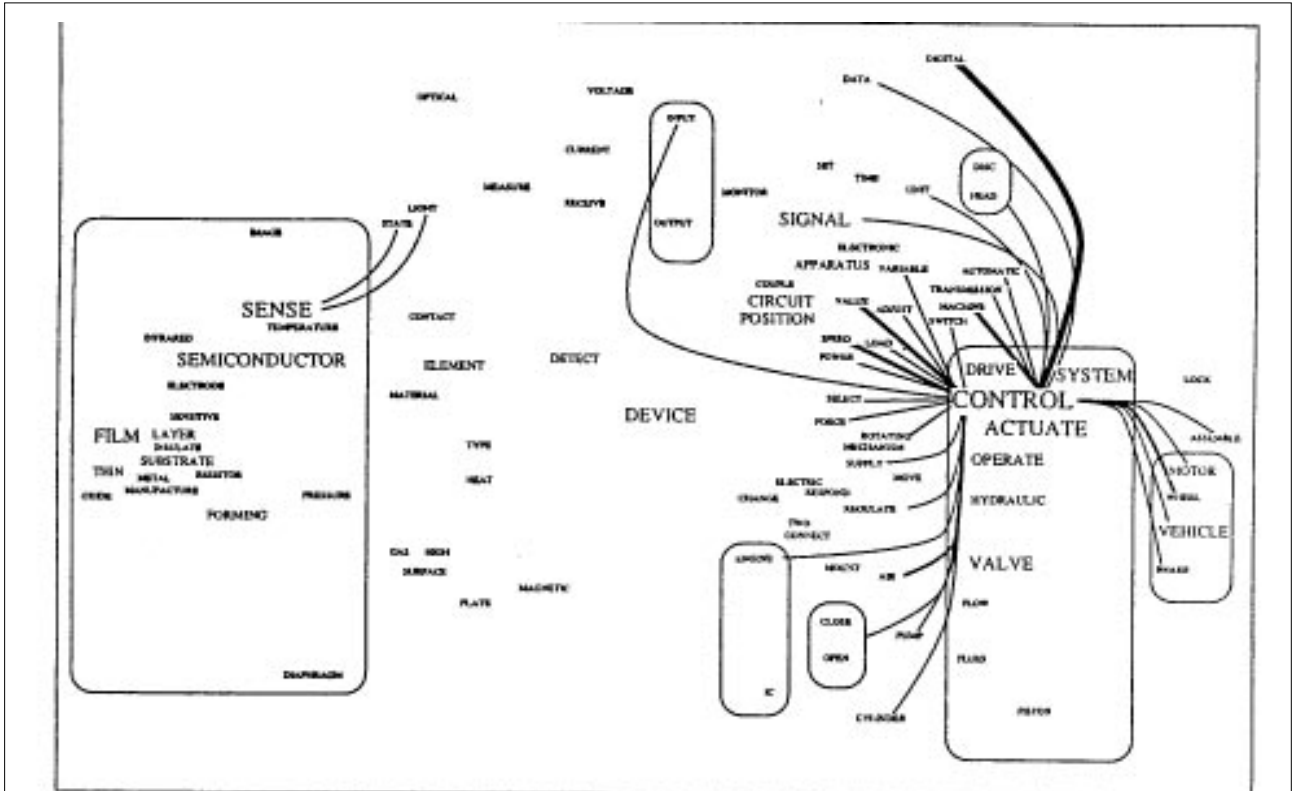
BAR : BARBAS C.F.	COL : COLLET T.A.	LER : LERNER R.A.	HOO : HOOGENBOOM H.R.
BRE : BRENNER S.	KER : KERR J.M.	PAV : PAVIA M.R.	NIK : NIKOLAIEV V.

Source : Données spécialement fournies à Y. Okubo par H. Small de l'Institute for Scientific Information.

Cette figure présente un exemple de co-citations emprunté à une étude de l'ISI consacrée à la “chimie combinatoire”, domaine d'avenir pour la mise au point de nouveaux médicaments, et à ses applications biologiques. Les abréviations se réfèrent à des chercheurs travaillant dans ce domaine. Les lignes reflètent les liens entre ces personnes, qui sont révélées par des co-citations. La dimension des points de référence reflète l'intensité de ces liens.

L'agrégat (*cluster*) des publications co-citées témoigne de l'existence, identifiée par des mots-clés associés (voir le chapitre 5), de plusieurs groupes de scientifiques partageant les mêmes intérêts et les mêmes références. En 1992, un premier agrégat s'est formé autour de l'un des auteurs, Barbas, qui se situe alors au centre de ces recherches. En 1993, l'agrégat grandit ; on trouve trois articles ajoutés autour de Kerr, et le sujet se déplace vers un autre sous-groupe des mêmes recherches (“les groupes péptidiques”, thème identifié par les titres des publications).

Exemple n° 24.
Carte du “champ de la technologie” dans l’optomécatronique



Source Noyons et Van Raan (1993).

Cette carte établie selon une technique matricielle, présente le côté “technologie” des travaux dans le domaine de l’optomécatronique en 1991. Elle est basée sur une analyse des fréquences de co-occurrences des mots-clés spécifiques à ce domaine dans des brevets de technologie. Les mots associés sont positionnés sur la carte en fonction de leurs relations mutuelles, y compris les “liens cachés” qui seraient difficiles à apercevoir sous d’autres formes de présentation. Les distances relatives entre des sujets de recherche définis par mots-clés (dont la taille sur la cartes indique la fréquence) reflètent leurs relations cognitives.

Exemple n° 25.

Matrice des collaborations internationales, 1981-86 : 98 pays et huit domaines scientifiques

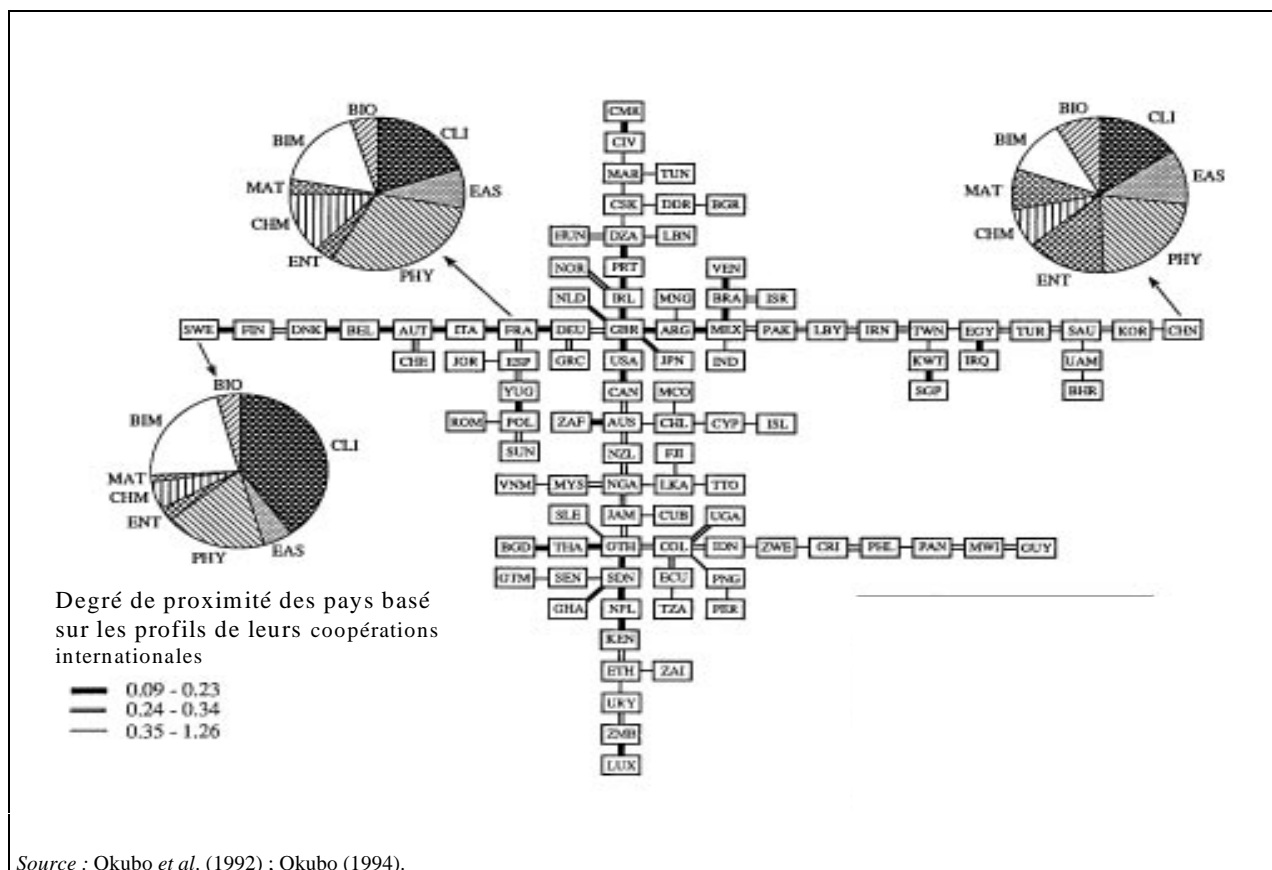
	MAT	PHY	CHM	ENT	EAS	BIO	BIM	CLI	Total
États-Unis	4 077	16 579	6 952	5 246	8 079	5 949	15 686	20 698	84 216
Royaume-Uni	1 168	7 334	4 099	1 860	3 601	2 492	6 639	9 158	36 763
Allemagne	962	9 964	3 540	1 446	2 935	1 790	5 687	5 714	32 219
France	931	8 387	3 433	898	2 263	1 280	4 530	5 287	27 521
Canada	1 335	3 376	2 103	1 614	2 119	2 001	3 501	4 967	21 129
Italie	410	4 70	1 699	402	1 268	511	2 160	3 344	14 652
Suisse	193	5 880	984	426	508	374	2 150	3 616	14 213
Monde	13 758	89 429	41 275	20 255	32 103	28 282	64 589	92 488	382 179

Source : Voir exemples 25A, B
 MAT (mathématiques), PHY (physique), CHM (chimie), ENT (ingénierie-technologie), EAS (Terre-espace), BIO (biologie fondamentale et appliquée), BIM (biomédicale), CLI (médecine clinique).

Les exemples 25 et 25A et B mettent en évidence les différents comportements des pays. Le tableau de l'exemple 25 présente le nombre de liens internationaux, mesurés en termes de co-signatures, pour les sept pays les plus importants dans la collaboration scientifique internationale, ainsi que le total mondial, le tout ventilé selon les huit principaux domaines scientifiques (pour les abréviations des noms des pays, voir l'annexe). Il s'agit donc là d'une analyse multidimensionnelle comportant 98 fois huit dimensions.

Sur la base de ce tableau on peut construire, pour chaque pays, un graphique en secteurs ("camembert") qui représente ses liens internationaux ventilés par domaine scientifique. Comme ces proportions varient selon les pays, chaque pays a son profil. Sur cette base, les pays sont groupés, dans les cartes, en agrégats (*clusters*) : ainsi, ceux qui se ressemblent s'assemblent.

Exemple n° 25A.
“L’arbre de longueur minimum”
Similitudes des pays basées sur le profil de leurs coopérations internationales, 1981-86

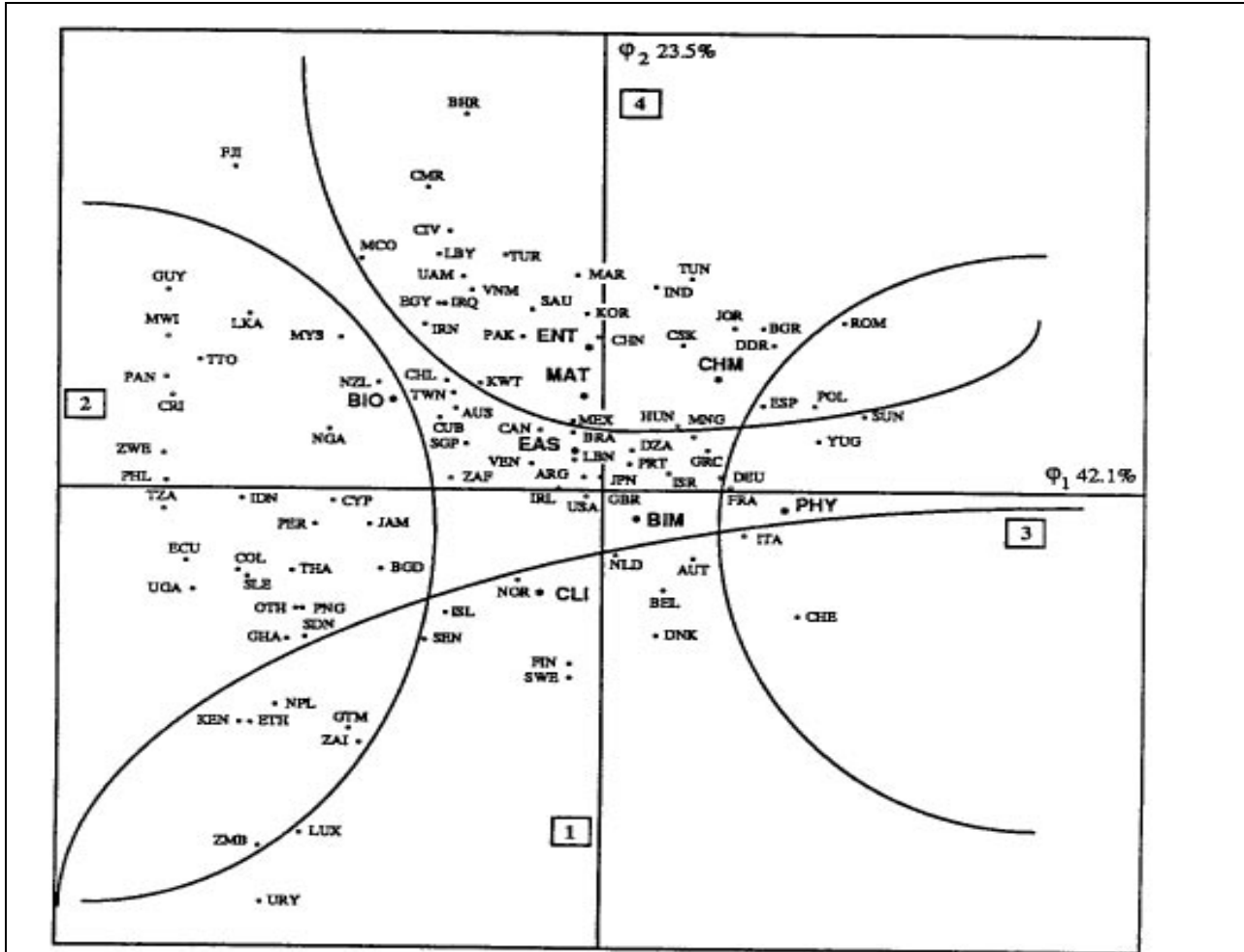


Cette “carte” est créée au moyen de la technique de classement dite de “l’arbre de longueur minimum” (ALM). Elle classe les 98 pays de l’exemple 25 par rapport au “modèle du monde” en fonction de la proximité des profils de leurs liens internationaux dans les huit domaines scientifiques au cours de la période. Chaque rectangle représente un pays et remplace, pour des raisons techniques évidentes, les “camemberts” correspondants. Les “degrés de proximité” entre pays sont illustrés par l’épaisseur des lignes indiquant les liens.

Trois autres “graphiques” illustrent, d’une part, les profils des deux “pays extrêmes” de la plus longue ligne horizontale (Suède et Chine) dont les liens de coopération sont fort différents, et, d’autre part, celui de la France. Les deux premiers sont assez éloignés du “modèle monde” qui, se trouvant près du centre de l’arbre, est fortement influencé par les structures de coopération des plus grands pays scientifiques (dont l’Allemagne, les États-Unis, la France, et le Royaume-Uni). De même, il y a de fortes différences dans les liens de coopération entre les pays de l’axe vertical, sur lequel prennent place les pays allant du Cameroun (en haut) au Luxembourg (en bas).

Tous les pays concentrent leurs échanges internationaux dans des champs spécifiques, quels que soient la taille du pays considéré et ses efforts pour adhérer au centre de gravité de l’activité scientifique. Sur la première voie principale, les pays européens (plutôt spécialisés en médecine clinique) s’alignent à gauche, jusqu’au Royaume-Uni (GBR), tandis que les pays arabes et islamiques ainsi que les pays d’Asie s’alignent à droite (tous plutôt orientés vers des coopérations dans les domaines de la physique et de l’ingénierie/technologie). On peut évoquer, à travers cet ALM, de nombreux facteurs historiques, culturels, économiques et politiques.

Exemple n° 25 B.
Positions relatives en coopération internationale des 98 pays
déterminées par les huit champs scientifiques, 1981-86



Analyse factorielle des correspondances

- (1) Pays ayant une forte activité en médecine clinique.
- (2) Pays ayant une forte activité en biologie fondamentale et appliquée.
- (3) Pays ayant une forte activité en physique.
- (4) Pays ayant une forte activité en ingénierie, mathématiques et chimie.

MAT (mathématiques), PHY (physique), CHM (chimie), ENT (ingénierie-technologie), EAS (Terre-espace), BIO (biologie fondamentale et appliqué), BIM (biomédicale), CLI (médecine clinique)

Les axes phi 1, 2, 3 représentent la totalité des informations contenues dans le tableau analysé. Les premiers axes (phi 1, 2) sont ceux qui portent l'essentiel des informations et représentent le mieux l'ensemble d'une matrice.

Source : Okubo *et al.* (1992) ; Okubo (1994).

Cette carte utilise l'analyse factorielle des correspondances pour "visualiser" la collaboration internationale et ses spécificités, toujours pour les 98 pays et les huit domaines de la science de l'exemple 25 en positionnant les éléments (ici les pays et les champs scientifiques).

Les "clés" permettant d'interpréter cette carte sont les suivantes :

- plus les pays sont rapprochés sur le diagramme, plus leurs profils de collaboration avec l'ensemble des pays dans les huit domaines scientifiques se ressemblent ;
- plus un pays est proche d'un domaine, plus sa collaboration dans ce domaine est importante, comparée à celle des autres pays ;
- plus un pays est proche du centre du diagramme, plus sa collaboration dans les divers domaines scientifiques est équilibrée.

Pour illustrer l'emploi de ces "clés", on peut prendre comme exemple le groupe situé près du centre de la carte. Ce groupe qui comprend, entre autres pays, le Japon (JPN) et les États-Unis (USA), a des modes de coopération semblables, avec une collaboration internationale à peu près équivalente dans chacun des domaines scientifiques.

Le groupe situé dans le cadre supérieur gauche (quadrant n°12), qui comprend, par exemple, l'Australie (AUS), le Canada (CAN) et la Nouvelle-Zélande (NZL), concentre sa collaboration dans les quatre domaines indiqués (BIO, EAS, MAT, ENT), la Nouvelle-Zélande ayant un degré de spécialisation élevé surtout en biologie.

Les pays nordiques - Islande (ISL), Norvège (NOR), Finlande (FIN) et Suède (SWE) - se concentrent sur la médecine clinique (CLI) (quadrant n° 1), tout en restant très proches du Danemark (DNK), de la Belgique (BEL) et de Pays-Bas (NLD) qui sont toutefois plus proches du domaine biomédical (BIM) du quadrant n° 3.

NOTES

1. Cette liste suit celle de Harada (1986, pp. 139-143), qui a cependant présenté six catégories au lieu des quatre décrites ici.
2. Elle est éditée à l'Information Science and Scientometrics Research Unit (ISSRU) de la bibliothèque de l'Académie des sciences de Hongrie (*Akadémiai Kiado és Nyomda Vallalat*), Budapest, Hongrie. Le premier numéro est paru en septembre 1978.
3. Pour les bases de données sur les brevets voir aussi OCDE, 1994, section 3.3.
4. *Physikalische Berichte*, une base de données produite par la société allemande *Deutsche Physikalische Gesellschaft*.
5. Par exemple, une étude comparative de Moed (1989) propose un état des mathématiques en Italie, mesuré de deux façons différentes. La base "fixed" indique un degré d'activité très inférieur à celui donné par l'ensemble variable. Ce décalage est dû à l'apparition de deux revues de mathématiques, *Bollettino della Unione Matematica Italiana* et *Annali di Matematica Pura et Applicata*, dans lesquelles les scientifiques italiens ont une forte présence et qui ont été incorporées dans le SCI pendant la période gelée de *fixed-journal set*.

RÉFÉRENCES

- Anderson, J., P.M.D. Collins, J. Irvine, P.A. Isard, B.R. Martin, F. Narin, et K. Stevens (1989), "On-line Approaches to Measuring National Scientific Output: A Cautionary Tale", *Science and Public Policy*, vol. 15(3), pp. 153-161.
- Bradford, S.C. (1950), *Documentation*, Public Affairs Press, Washington DC.
- Braun, T., W. Glänzel, et A. Schuber (1991), "The Bibliometric Assessment of UK Scientific Performance: Some Comments on Martin's "Reply", *Scientometrics*, vol. 20(2), pp. 359-362.
- Bush, V. (1960), "Science and the Public Welfare", dans *Science: The Endless Frontier*, National Science Foundation (Reprint), Washington, DC, pp. 17-27.
- Callon, M., J.P. Courtial et H. Penan (1993), *La scientométrie*, Presses universitaires de France, Paris.
- Callon, M. et L. Leydesdorff (1987), "La recherche française est-elle en bonne santé ?", *La Recherche*, vol. 18(186), pp. 412-419.
- Chauvin, R. (1991), "Évaluer, évaluer", *La Recherche*, vol. 22, p. 782.
- Chelimsky, E. (1991), "On the Social Science Contribution to Governmental Decision-Making", *Science*, vol. 254, pp. 226-231.
- Cole, J. et N.B. Eales (1917), "The History of Comparative Anatomy: A Statistical Analysis of the Literature", *Science Progress*, vol. 11(4), pp. 578-596.
- Cunningham, E.R. (1935), "The Present Status of the Publication of Literature in the Medical and Biological Sciences", *Medical Library Association Bulletin*, vol. 24, pp. 64-81.
- Frame, J.D. (1977), "Mainstream Research in Latin America and the Caribbean", *Interciencia*, vol. 2(3), pp. 143-148.
- Garfield, E. (1968), *World Brain or Memex? Mechanical and Intellectual Requirements for Universal Bibliographic Control, The Foundations of Access to Knowledge*, Syracuse University Press, New York.
- Garfield, E. (1972), "The Design and Production of a Citation Index", dans *Citation Indexing-Its Theory and Application in Science, Technology and Humanities*, John Wiley & Sons, New York, pp. 19-36.
- Garfield, E. (1975), "Introduction", dans *Journal Citation Reports: Science Citation Index*, vol. 9(3).
- Garfield, E. (1979a), "How Do We Select Journals for Current Contents?", *Current Contents*, 5 novembre.
- Garfield, E. (1979b), "Is Citation Analysis a Legitimate Evaluation Tool?", *Scientometrics*, vol. 1, pp. 359-375.
- Garfield, E. (1988), "French Research: Citation Analysis Indicates Trends are More than Just a Slip of the Tongue", *Current Contents*, vol. 23, p. 9.
- Glänzel, W. et U. Schoepflin (1994), "Little Scientometrics, Big Scientometrics . . . and Beyond", *Scientometrics*, vol. 30(2-3), p. 375.
- Harada, M. (1986), "La tendance de la bibliométrie", *Operations Research*, mars, pp. 139-143.
- Hirasawa, R. (1995), "Integrative Management of Technology - Analysis and Its Implications", R&D Dynamics Network Meeting in Stockholm, 7-9 novembre.
- Hulme, E. W. (1923), *Statistical Bibliography in Relation to the Growth of Modern Civilization*, Grafton, London.
- Institute for Scientific Information (1981) *Science Citation Index 1981 Guide & Lists of Source Publications*, Institute for Scientific Information, Inc., Philadelphia.
- Irvine, J. et B.R. Martin (1980), "The Economic Effects of Big Science: The Case of Radio Astronomy", Science & Technology Indicators Conference, OCDE, Paris.
- Katz, J.S. et D.M. Hicks, A Systemic View of British Science, *Scientometrics* 35, n°1 (1996), pp. 133-154.

- Kealey, T. (1991), "Government-Funded Academic Science is a Consumer Good, Not a Producer Good: A Comparative Reassessment of Britain's Scientific and Technological Achievements since 1794 and a Comment on the Bibliometry of B. Martin and J. Irvine", *Scientometrics*, vol. 20(2), pp. 369-394.
- Kobayashi, S. (1987), "Internationalisation of Japanese Journals and their Contribution to the Basic Science", *Daigaku Kenkyu*, vol. 1, pp. 57-76.
- Leven, O. (1982), *Report of the National Survey Committee on Biochemistry*, Staatsuitgeverij, The Hague.
- Leydesdorff, L. (1991), "On the 'Scientometric Decline' of British Science. One Additional Graph in Response to Ben Martin", *Scientometrics*, vol. 20(2), pp. 363-368.
- Leydesdorff, L. (1995), *The Challenge of Scientometrics : The development, measurement, and self-organisation of scientific communication*, DSWO Press, Leyde.
- Lotka, A.J. (1926), "The Frequency Distribution of Scientific Productivity", *Journal of the Washington Academy of Science*, vol. 16(12), pp. 317-323.
- Magri, M.H. et A. Solari (1996), "The SCI Journal Citation Reports: A Potential Tool for Studying Journals?", *Scientometrics*, vol. 35, pp. 93-117.
- Martin, B.R. (1991), "The Bibliometric Assessment of UK Scientific Performance. A Reply to Braun, Glänzel and Schubert", *Scientometrics*, vol. 20(2), pp. 333-357.
- Martin, B.R. et J. Irvine (1984), "CERN: Past Performance and Future Prospects", *Research Policy*, vol. 13, pp. 311-342.
- Martin, B.R. et J. Irvine (1985) "Basic Research in the East and West: A Comparison of the Scientific Performance of High-Energy Physics Accelerators", *Social Studies of Science*, vol. 15, pp. 293-394.
- Merton, R. K. (1957b), "Social and Democratic Social Structure", dans *Social Theory and Social Structure*, Free Press, New York, pp. 550-561.
- Merton, R.K. (1957a), "Priorities in Scientific Discovery", *American Sociological Review*, vol. 22, p. 635.
- Ministère de l'Éducation, de la Science et de la Culture (1987), "Étude comparative internationale du nombre d'articles scientifiques", Japon.
- Miquel, J.F. et J.C. Doré (1981), "Research on New Drugs: A System Analysis Approach", dans *Drugs of Today XVII (11)*, pp. 487-501.
- Miquel, J.F. et Y. Okubo (1990), "Indicators to Measure Internationalization of Science", *Consequences of the Technology Economy Programme for the Development of Indicators*, OCDE, Paris.
- Miquel, J.F. et Y. Okubo (1994), "Structure of International Collaboration in Science-Part II: Comparisons of Profiles in Countries using a Link Indicator", *Scientometrics* 29 (2), pp.271-297.
- Moed, H.F. (1989), "New Developments in Online Bibliometric Analysis", dans *Science and Technology Indicators*, DSWO Press, Leiden, pp. 129-146.
- Moed, H.F., W.J.M. Burger, J.G. Frankfort, et A.F.J. Van Raan (1983), "On the Measurement of Research Performance: The Use of Bibliometric Indicators", Université de Leyde, Leyde.
- National Science Board (1989), "Public Attitude towards Science in General", dans *Science and Engineering Indicators 1989*, National Science Foundation, Washington, DC.
- National Science Board (1989), *Science and Engineering Indicators 1989*, National Science Foundation, Washington, DC.
- National Science Foundation (1972), *Science and Engineering Indicators 1972*, National Science Foundation, Washington, DC.
- National Science Foundation (1989), *Science & Engineering Indicators 1989*, National Science Foundation, Washington, DC.
- Nederlands Observatorium van Wetenschap en Technologie (Centrum voor Wetenschaps-en Technologie-Studies (1994), *Wetenschaps-en Technologie-Indicatoren 1994*, Rijksuniversiteit, Leyde.

- Niwa, F. et H. Tomizawa (1995), "Development of General Indicators of Science and Technology", dans *Proceedings of the Fourth International Conference on Science and Technology Indicators*, Antwerp.
- Noyons, E.C.M. et A.F.J. Van Raan (1993), "Science Base of Technology - Bibliometric Mapping as a Tool for National Science and Technology Policy : Part II: Optomechatronics", *Rapport CWTS 93-07*, Université de Leyde.
- Noyons, E.C.M. et A.F.J. Van Raan (1995), "Mapping the Development of Neural Network Research: Structuring the Dynamics of Neural Network Research and an Estimation of the German Activity", Report CWTS 95-06, Centrum voor Wetenschappen-en Technologie-Studies.
- OCDE (1989), "Les analyses bibliométriques", dans *La mesure de l'output de R-D, La mesure des activités scientifiques et techniques, statistiques de R-D et mesure des résultats dans l'enseignement supérieur*, complément du "Manuel de Frascati", Paris.
- OCDE (1994), "Les données sur les brevets d'invention et leur utilisation comme indicateurs de la science et de la technologie - Manuel Brevet 1994", Paris.
- Okubo, Y., J.F. Miquel, L. Frigoletto et J.C. Doré (1992), "Structure of International Collaboration in Science; Typology of Countries through Multivariate Techniques using a Link Indicator", *Scientometrics* 25, vol.2, pp. 321-351.
- Okubo, Y. (1993), "International Co-operative Network in Drug Innovation International R&D Network", Conference de Kyoto, mai 1993.
- Okubo, Y. (1994), L'Internationalisation de la Science : Création d'indicateurs bibliométriques pour une mise à jour de l'activité scientifique internationale du Japon, thèse de Doctorat.
- Observatoire des Sciences et des Techniques (1992), "Science & Technologie Indicateurs", *Economica*, Paris.
- Otsu, K. (1983), "A Bibliometric Study of Japanese Science and Social Science Publications", *Library and Information Science*, vol. 21, pp. 19-27.
- Ourisson, G. (1991), "Une évaluation pertinente", *La Recherche*, vol. 22, pp. 1209-1210.
- Pendlebury, D. (1991), "Research Papers: Who's Uncited Now?", *Science*, vol. 251, p. 25.
- Price, D. (1963), *Little Science, Big Science*, Columbia University Press, New York.
- Price, D. (1964), *The Science of Science*, London Souvenir Press, London.
- Price, D. (1965), "Networks of Scientific Papers", *Science*, vol. 149, pp. 510-515.
- Price, D. (1969), "Policies for Science?", *Melbourne Journal of Politics*, vol. 2, pp. 1-8.
- Pritchard, J. (1969), "Statistical-Bibliography or Bibliometrics?", *Journal of Documentation*, vol. 25(4), pp. 348-349.
- Rappa, M.A. (1989), "Bibliometric Methods for Monitoring the Emergence of New Technologies", Rapport n° WP3049-89-BPS, Massachusetts Institute of Technology, Massachusetts.
- Royal Society (1986), *Evaluation of National Performance in Basic Research*, étude effectuée par le service de recherches de la Royal Society pour l'Advisory Board for the Research Councils, The Royal Society, Londres.
- Seglen, Per O. (1992), "The Skewness of Science", *Journal of American Society for Information Science*, vol. 43, pp. 628-638.
- Sivertsen, G. (1991), "Should a New Database for International Comparison be More Restricted in Journal Coverage?", dans *Science and Technology in a Policy Context*, Actes du Colloque CE-Leyde sur les indicateurs de science et technologie, DSWO Press, Leyde.
- The Economist* (1992), "Pity the Typists", 18 janvier, p. 87.
- Van Raan, A.F.J. et R.J.W. Tijssen (1990), "An Overview of Quantitative Science and Technology Indicators Based on Bibliometric Methods", Technology Economy Programme for the Development of Indicators, OCDE, Paris.
- Wade, N. (1975), "Citation Analysis: A New Tool for Science Administrators", *Science*, vol. 188, pp. 429-432.

**ANNÈXE I.
INDÈXE DES CODES ISO**

ARG	Argentine	KWT	Koweït
AUS	Australie	LBN	Liban
AUT	Autriche	LBY	Libyenne, Jamahiriya arabe
BEL	Belgique	LKS	Sri Lanka
BGC	Bangladesh	LUX	Luxembourg
BGR	Bulgarie	MAR	Maroc
BHR	Bahréïn	MCO	Monaco
BRA	Brésil	MEX	Mexique
CAN	Canada	MLI	Mali
CHE	Suisse	MNG	Mongolie
CHL	Chili	MWI	Malawi
CHN	Chine	MYS	Malaisie
CIV	Côte d'Ivoire	NGA	Nigéria
CMR	Cameroun	NLD	Pays-Bas
COL	Colombie	NOR	Norvège
CRI	Costa Rica	NPL	Népal
CSK	Tchécoslovaquie	NZL	Nouvelle-Zélande
CUB	Cuba	PAK	Pakistan
CYP	Chypre	PAN	Panama
DDR	Allemande, Rép démocratique	PER	Perou
DEU	Allemagne, Rép fédérale	PHL	Philippines
DNK	Danemark	PNG	Papouasie-Nouvelle-Guinée
DZA	Algérie	POL	Pologne
ECU	Equateur	PRT	Portugal
EGY	Egypte	ROM	Roumanie
ESP	Espagne	SAU	Arabie saoudite
ETH	Ethiopie	SDN	Soudan
FIN	Finlande	SEN	Sénégal
FJI	Fidji	SGP	Singapour
FRA	France	SLE	Sierra Leone
GBR	Royaume-Uni	SUN	URSS
GHA	Ghana	SWE	Suède
GRC	Grèce	THA	Thaïlande
GTM	Guatemala	TTO	Trinité-et-Tobago
GUY	Guyana	TUN	Tunisie
HUN	Hongrie	TUR	Turquie
IDN	Indonésie	TWN	Taïwan, Province de Chine
IND	Inde	TZA	Tanzanie, République Unie de
IRN	Iran	UGA	Ouganda
IRL	Irlande	URY	Uruguay
IRQ	Iraq	UAM	Emirats Arabes
ISL	Islande	USA	Etats Unis
ISR	Israël	VEN	Venezuela
ITA	Italie	VNM	Viet Nam
JAM	Jamaïque	YUG	Yougoslavie, Rép. Socialiste Fédérative de
JOR	Jordanie	ZAF	Afrique du Sud, Rép. d'
JPN	Japon	ZAI	Zaïre
KEN	Kenya	ZMB	Zambie, République de
KOR	Corée, République de	ZWE	Zimbabwe

STI WORKING PAPERS

1996

1. Embodied Technology Diffusion: An Empirical Analysis for 10 OECD Countries
George Papaconstantinou, Norihisa Sakurai and Andrew Wyckoff
2. The Impact of R&D and Technology Diffusion on Productivity Growth: Evidence for 10 OECD Countries in the 1970s and 1980s
Norihisa Sakurai, Evangelos Ioannidis and George Papaconstantinou
3. Short-term Indicators: Using Qualitative Indicators to Update Production Indices
Paul Schreyer and Corinne Emery
4. SMEs and Employment Creation: Overview of Selected Quantitative Studies in OECD Member Countries
Paul Schreyer
5. Globalisation and Competitiveness: Relevant Indicators
Thomas Hatzichronoglou
6. Factors Influencing the Steel Work Force: 1990 to 1995
Donald F. Barnett
7. Measuring R&D in the Services
Alison Young
8. The Evolution of Skills in OECD Countries and the Role of Technology
A. Colecchia and G. Papaconstantinou

1997

1. Indicateurs bibliométriques et analyse des systèmes de recherche : méthodes et exemples
Yoshiko Okubo