

SUR L'USAGE RECENT DES INDICATEURS BIBLIOMETRIQUES COMME OUTIL D'EVALUATION DE LA RECHERCHE SCIENTIFIQUE.

par

Laurence Coutrot

(CMH-CNRS, laurence.coutrot@ens.fr)

Bulletin de Méthodologie sociologique (BMS), N°100, octobre 2008.

Pour le meilleur et pour le pire, les grandes manœuvres sont lancées dans la recherche française sur le front de la bibliométrie : celle-ci cesse désormais d'être un outil réservé aux documentalistes et spécialistes de l'information. Les princes qui nous gouvernent s'en sont emparé à des fins d'évaluation des disciplines, des laboratoires, des individus.

Les personnels de recherche doivent pouvoir s'informer, réfléchir et critiquer, et, éventuellement inventer des usages de la bibliométrie qui en limiterait les possibles effets pervers.

L'affaire n'est pas neuve : il y a vingt ans déjà, à la Direction du CNRS, les activités de construction d'indicateurs scientifiques étaient à la mode. En 2004, une tentative de classement des revues par le Département SHS du CNRS avait provoqué quelques réactions¹. Mais cette fois, l'affaire est sérieuse : dans le document de « caractérisation des UMR » pour la vague D de contractualisation du quadriennal, le « facteur h » figure en bonne place, aux côtés de l'impact moyen de la discipline, etc. On est passé du classement des revues à celui des individus, ce qui est toute autre chose.

Cette courte note n'a évidemment pas pour fonction de balayer l'ensemble du problème de la bibliométrie. Je vais seulement tenter d'apporter quelques informations essentielles concernant les principes et les outils de la bibliométrie et de pointer quelques problèmes liés à son application à des fins d'évaluation de la recherche.

A l'initiative de « Go!doc », réseau de documentalistes CNRS de la région parisienne (<http://www.godoc.u-psud.fr/>), et avec le concours du service de formation permanente du CNRS, une réunion d'information sur les indicateurs bibliométriques a eu lieu le mardi 23 septembre 2008. L'amphithéâtre était plein : le sujet attire les foules, ITA et CNRS de toutes disciplines. La séance était présidée par Rémi Barré, professeur au CNAM, ancien président de l'OST (Observatoire des sciences et des techniques) et actif à la DGRI au Ministère de la recherche. On pouvait craindre que cette séance fût uniquement un lieu de diffusion de la vulgate technocratique sur le sujet, cela ne s'est pas vérifié : il y eut des exposés critiques et un vrai débat, d'autres suivront sans aucun doute.

La séance débute par une présentation de l'un des outils disponibles pour analyser les périodiques scientifiques, le Web of Science : « *Emergence et construction des outils bibliométriques dans le Web of Science, répertoire des outils et leur justification* ». L'exposé est fait par Guillaume Rivalle et Nicolas Espeche, tous deux salariés de Thomson Reuters,

¹ « Pour une Realpolitik de la recherche » au sujet de l'enquête du CNRS sur les périodiques aidés par le département des sciences sociales et de la société, *Actes de la recherche en sciences sociales*, 2004/5, n. 155, pp.101-104.

compagnie qui a repris l'Institute for Scientific Information (ISI), créé dans les années 1950 par E. Garfield. Le Web of Science (WoS) et le Web of Knowledge, entre autres, c'est eux. Le logiciel documentaire très répandu, « Endnote », c'est également eux. L'entreprise est basée à New York et Philadelphie, c'est très américain et c'est « Big Money ». Ils recensent à l'heure actuelle environ 10.000 revues scientifiques (dont 1.865 pour le Social Science Citation Index) dans 42 langues différentes. Ils emploient plus de 50.000 salariés dans le monde entier. Ces deux présentateurs sont des commerciaux excellents : ils vantent les mérites de leur outil. Le site de Thomson Reuters vaut la visite².

Pour faire vite, on part du principe que la publication est une trace essentielle de l'activité scientifique et que la « notoriété », le fait d'être cité, est un indicateur de « l'influence » d'un auteur. Qui cite qui, qui est cité par qui et combien de fois... On calcule ainsi un « facteur d'impact » d'un auteur ou d'un groupe d'auteurs, d'un laboratoire, d'une institution, etc.

Les revues recensées sont en principe sélectionnées en fonction de critères de « sérieux ». Chaque article recensé fait l'objet d'une indexation (par mots-clé), les noms des auteurs sont répertoriés par leur institution de rattachement. Là, les problèmes commencent : l'institution est celle mentionnée par la revue et cette information est très lacunaire, en France et en SHS en particulier. D'où la volonté de n'autoriser que deux institutions de rattachement par chercheur affichée dans le projet « Normadresse ».

Les principaux outils (accessibles via le portail de l'INIST pour les agents CNRS) sont les suivants : Search History, Analyse Research, Journal Citation report (JCR), Cited reference search. En outre, le RSG (Research Service Group) offre des traitements personnalisés, payants. Le JCR fournit un certain nombre d'information sur les revues regroupées en catégories : nom du périodique, nombre d'articles publiés pour une année donnée, nombre de fois où un article a été cité (par année ou cumulé), liste des périodiques qui ont cité un article de ce périodique pour une année de référence (*Cited search*). On peut ainsi retrouver des articles cités dans des revues non incorporées dans la base de données du WoS. Malheureusement, les bugs sont nombreux : on trouve ainsi dans la liste des revues qui citent ARSS (*Actes de la recherche en sciences sociales* « Euvres complètes », et « Condition Lit Double », sic...).

Toutes ces informations permettent de déterminer le « facteur d'impact » du périodique (« Journal Impact Factor »), calculé à partir du quotient : nombre de fois où, en 2007, les articles du périodique X publiés en 2005 et 2006 sont cités dans un ensemble de périodiques de la catégorie, rapporté au nombre d'articles publiés en 2005 et 2006 dans le périodique X. Ainsi, les 77 articles de *l'American Journal of Sociology* publiés en 2005 et 2006 sont réputés avoir été cités 257 fois en 2007, ce qui donne un facteur d'impact de $257/77 = 3,338$. On peut aussi regarder les choses sous un autre angle (*Citing*) et regarder dans quelles revues sont publiés les articles cités dans les articles d'un périodique donné pour une année de référence.

Un premier coup d'œil à cette base montre qu'elle analyse 96 revues sous la rubrique « sociologie », 57 sous « Sciences sociales et interdisciplinaires » et que, pour être précis, il faudrait aussi considérer d'autres catégories, notamment l'anthropologie (ainsi la *Revue française de sociologie* est répertoriée dans la première catégorie, mais *Actes de la Recherche*

² <http://scientific.thomsonreuters.com/products/wos/>

en sciences sociales dans la seconde). L'immense majorité de ces revues sont américaines, il y a quelques grandes revues françaises et allemandes, mais peu d'italiennes ou d'espagnoles. En gros, tout ce qui n'est pas anglophone est largement sous-représenté et souffre de bugs évidents (les umlaut et autres accents aigus, c cédille, etc. sont royalement ignorés).

Un second exposé portait sur « *Indicateurs de positionnement institutionnel pour l'aide à la décision* » : comment les indicateurs bibliométriques sont à l'heure actuelle utilisés pour guider la politique scientifique (G. Filiatreau, Directrice de l'Observatoire des sciences et des techniques, OST). Il a également été question du projet « Normadresse », d'où il ressort que la politique de « normalisation des adresses » est en fait un levier pour pouvoir appliquer la bibliométrie à des fins de pilotage de la recherche. Chercheurs et « publiants », l'ordre dans lequel vous mentionnez vos institutions de rattachement est tout sauf trivial. Par exemple, modifier l'ordre d'une publication à l'autre, c'est s'exposer à ne pas être recensé dans la base bibliométrique.

Les deux premiers exposés de l'après-midi ont été faits par des chercheurs : Yves Langevin, Président de la section 17 du Comité national (Sciences de la Terre et Univers lointains) a exposé comment en commission 17 les indicateurs bibliométriques ont été utilisés, aux côtés d'autres critères, pour le concours DR2.

Deux points importants : Y. Langevin montre comment la commission a construit une batterie d'indicateurs extrêmement complexes (et non pas un simple « facteur h » brandi comme une hache) : pas moins de 14 indicateurs bibliométriques ont été utilisés. Si cette batterie d'indicateurs permet, en gros, de confirmer si les candidats se trouvent ou non dans la première moitié du classement final, en revanche la corrélation avec les positions dans les premiers rangs n'est pas bonne. D'autres indicateurs (jugements par les pairs, positions de responsabilité dans le milieu scientifique, etc.) sont de poids quand il s'agit de déterminer les positions dans les premiers rangs.

Un détail amusant : en section 17, toutes les revues scientifiques sont considérées comme étant de rang A.

Y. Langevin souligne que l'utilisation de la bibliométrie pour l'évaluation pénalise les chercheurs qui ont une mobilité thématique (puisque'il faut plusieurs années pour se refaire une réputation à l'intérieur d'un groupe thématique donné).

Enfin, ce système entraîne une série d'effets pervers dont l'uniformisation des pratiques scientifiques, la multiplication des publications avec tronçonnement des résultats, etc. Comme Y. Langevin est optimiste, il prédit qu'à terme ce système d'évaluation par la bibliométrie s'autodétruit par uniformisation des pratiques.

Luc Segoufin, chercheur et membre de la commission d'évaluation de l'INRIA a montré sur un exemple précis, comment si l'on prend des bases différentes (Google scholar, Web of Science, Scopus, Citeseer, DBLP), on obtient des mesures d'impact différentes pour un même chercheur. C'est irréfutable : tout index a un biais, le mode de calcul des index est déterminant, ce n'est pas pertinent pour les évaluations interdomaines. Enfin, le coût des outils payants est lourd (il semblerait qu'on ait du mal à savoir le montant exact des contrats passés entre le CNRS et ses partenaires universitaires avec Thomson Reuters...).

L. Segoufin a repris dans son exposé les grandes lignes du rapport élaboré par la commission d'évaluation de l'INRIA sur le sujet. Voir <http://www.inria.fr/inria/organigramme/ce.fr.html> puis cliquer sur « indicateurs bibliométriques ».

On peut y lire en conclusion : « *Si les indicateurs peuvent donner des tendances sur un nombre réduit d'aspects de la vie scientifique, il convient d'être très circonspect dans leur usage en raison de la possibilité d'interprétations erronées, des erreurs de mesure (souvent considérables) et des biais dont ils sont affectés. Un usage abusif des indicateurs est facilité par la nature chiffrée du résultat qui introduit la possibilité d'établir dans l'urgence toutes sortes de statistiques, sans se préoccuper d'en analyser la qualité et le contenu, et en occultant l'examen d'autres éléments de la vie scientifique comme, par exemple, l'innovation et le transfert intellectuel et industriel.* »

Suivent un certain nombre de recommandations qui invitent à la plus grande prudence dans le maniement d'informations dont le caractère lacunaire et entaché d'erreurs est constamment souligné. Dans son exposé L. Segoufin conclut en montrant que ces indicateurs bibliométriques peuvent éventuellement servir à dégager « le haut de la pyramide », mais que le jugement par les pairs reste l'outil le plus sûr et, paradoxalement, le moins subjectif...

Cette réunion a été utile, un certain nombre d'objections ont pu être exprimées librement, reflétant souvent l'inquiétude par rapport à cette question : la notoriété peut-elle être prise pour un bon indicateur de la qualité scientifique ?

Un dernier point : il faut se garder de croire que les réticences à l'égard de la bibliométrie comme outil d'évaluation concernent les seules Sciences de l'Homme et de la Société. Cette réunion le montrait clairement, rejoignant ainsi les réflexions que l'on trouve dans cet excellent rapport produit par un groupe de mathématiciens américains³. Vous trouverez là un rappel des principales notions de bibliométrie, y compris le mode de calcul du « facteur h »⁴. Une lecture complémentaire nous est fournie par un chercheur canadien, Yves Gingras⁵. On y trouve un exposé clair des principes de calcul des indicateurs, du « facteur h » et du classement de Shangaï.

Les grandes manœuvres actuelles autour de la bibliométrie me rappellent étrangement les querelles passées autour de la mesure de l'intelligence par le « Quotient intellectuel » : mesurer rassure, mais l'on ne mesure pas ce que l'on croyait ou voulait mesurer.

En tout état de cause, il faut se garder d'oublier que les articles cités sont le résultat de filtrages successifs : ils ne sont que le dernier niveau, étroit, d'une fusée à plusieurs étages : les chercheurs produisent d'abord des communications ou des rapports, puis rédigent des articles, qu'ils soumettent à des revues ; une fraction variable de ces manuscrits arrive à publication, un sur dix ou un sur vingt, parfois moins. Enfin, un article publié peut, par une

³ Texte intégral : <http://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf> Résumé disponible sur <http://www.sauvonslarecherche.fr/spip.php?article2004>

⁴ Voir Hirsch J.E. (2005), An Index to quantify an individual's scientific output, *Proceedings of the National Academy of Sciences*, vol. 102, n. 46, pp. 16569-16572.

⁵ Gingras Y., (2008) La fièvre de l'évaluation de la recherche. Du mauvais usage de faux indicateurs. CIRST, UQAM. <http://www.cirst.uqam.ca> Voir ci-dessous l'« Introduction » de ce même rapport.

série de processus sociaux que l'on connaît fort mal⁶, parvenir à être cité. La citation est une mesure raisonnable de la visibilité d'un résultat scientifique. Quand bien même l'on voudrait départager les chercheurs qui travaillent de ceux qui « ne travaillent pas » (à supposer que cette expression ait un sens), le critère de citation est une mesure bien rustique. Sous des parures d'objectivité, cet indicateur réintroduit en fait ce qu'il prétendait éliminer ; c'est-à-dire, le jugement par les pairs.

Il est urgent pour la communauté scientifique en SHS de s'informer et d'élaborer une réflexion collective sur le sujet : que nous le voulions ou non, pour de bonnes ou de mauvaises raisons, ces outils bibliométriques sont déjà utilisés⁷. Certains de ces usages peuvent être relativement raisonnables, mais ils ne sont ni neutres ni sans effet de retours sur les pratiques de recherche. D'autres usages de ces indicateurs sont illégitimes. Il reste à obtenir que des outils qui ont quelque intérêt ne soient pas utilisés en dépit du sens commun.

⁶ Il y a vingt ans, Suzanne Cozzens à la NSF à Washington s'était déjà penchée sur le problème : Cozzens S. (1989), What do citations count? *Scientometrics*, vol. 15, n. 5-6, pp. 437-447.