La manufacture de l'évaluation scientifique : algorithmes, jeux de données, outils bibliométriques

David Pontille, Didier Torny

▶ **To cite this version:**

David Pontille, Didier Torny. La manufacture de l'évaluation scientifique : algorithmes, jeux de données, outils bibliométriques. Réseaux : communication, technologie, société, 2013, 177, pp.25-61. 10.3917/res.177.0025 . hal-00821956

HAL Id: hal-00821956

https://minesparis-psl.hal.science/hal-00821956

Submitted on 5 Jul 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La manufacture de l'évaluation scientifique : algorithmes, jeux de données et outils bibliométriques

David Pontille
Centre de Sociologie de l'Innovation
CNRS (UMR7185) - Mines-ParisTech
david.pontille@mines-paris.fr

&

Didier Torny[*]
Risques, Travail, Marchés, État
RiTME - INRA (UR 1323)
torny@ivry.inra.fr

Réseaux
2013, n°177, p. 25-61

[*] Les deux auteurs ont conventionnellement contribué de manière égale à la collecte et au traitement des données, à l'analyse et à l'écriture de l'article, aucune mesure techniquement fondée ne permettant de distinguer la contribution respective de chacun.

Résumé

Contre la vision d'une bibliométrie considérée comme un tout unifié et cohérent, cet article distingue trois composants majeurs des infrastructures informationnelles qui soutiennent l'évaluation de la recherche scientifique : les algorithmes, les jeux de données, et les outils bibliométriques. En retraçant la genèse et le succès de certains de ces composants, de 1960 à nos jours, les auteurs identifient plusieurs configurations dans lesquelles s'articulent de manière différentielle la position des producteurs de jeux de données, la cible privilégiée des algorithmes et la forme des outils bibliométriques. Cette histoire montre des croisements répétés entre bibliométrie et webométrie, et rappelle que toute évaluation s'appuie nécessairement sur une définition des pairs jugés pertinents.

Abstract

Against the vision of bibliometrics as a unified and coherent whole, this article identifies three major components of information infrastructures that support the evaluation of scientific research: algorithms, data sets, and bibliometric tools. Tracing the genesis and success of some of these components, from 1960 to the present day, the authors identify several configurations which differentially articulate the position of datasets producers, the target of algorithms, and the form of bibliometric tools. This diachronic narrative shows repeated crosses between bibliometrics and webometrics and recalls that any assessment is necessarily based on a definition of relevant peers.

> “A politics of formalism rests on several things. First, a formal representation is an abstraction: It takes away properties from a particular situation. Second, it is a simplification: It reduces the complexity of real life situations in order to make them formally (usually, but not exclusively, mathematically) tractable. Third, and most important, every formal representation contains choices about what to keep in (what is important) and what to throw out. All such choices are political”. (Star 1995, p. 113)[2]

Depuis plusieurs années, l’intensification de la compétition internationale pour l’excellence scientifique s’accompagne d’une prolifération des instruments de mesure destinés à évaluer systématiquement les universités, à rationaliser les organismes de recherche et à améliorer leurs performances. Si ce mouvement fait l’objet d’un véritable engouement de la part des gouvernements et des agences publiques dans plusieurs pays (Auranen et Nieminen 2010), il suscite également de nombreux débats, conduit à des prises de position critiques, et soulève des inquiétudes récurrentes sur les risques de la “quantophrénie” tous azimuts. Le phénomène semble bien identifié : la plupart des acteurs discutent et se disputent à propos des bienfaits et des dérives de la bibliométrie. Pour autant, on ne sait finalement pas vraiment de quoi sont faites les mesures quantitatives de la science.

Dans cet article, nous ferons volontairement un pas de côté vis-à-vis des conceptions qui font de “la” bibliométrie un tout unifié et cohérent. En nous appuyant sur la production scientométrique elle-même (articles de recherche, rapports) et sur les rares travaux portant sur l’histoire de la bibliométrie, nous proposons de revisiter en partie ses fondements et de déplier la grande hétérogénéité de ses éléments. Nous focaliserons en particulier notre attention sur les technologies de calcul qui équipent aujourd’hui les jugements portés sur les activités de recherche des institutions, des laboratoires, des équipes et des chercheurs individuels. Ce faisant, notre démarche prolonge la perspective initiée par B. Latour (1989) qui, avec la notion de “centre de calcul”, redonnait au formalisme ses ancrages matériels : loin d’être des opérations seulement représentationnelles, les calculs reposent au contraire sur une distribution entre une pluralité d’acteurs et d’instruments tangibles. G.C. Bowker et S.L. Star (1999) ont par la suite bien montré que les opérations de dénombrement, de codage et de classement s’accompagnent de la fabrique d’infrastructures informationnelles qui, à la manière des infrastructures physiques (e.g. réseau d’eau potable, ferroviaire, électrique…), résultent de décisions techniques distribuées entre plusieurs acteurs. Une fois élaborées, ces infrastructures informationnelles deviennent, pour une bonne part, invisibles, leur évidence partagée conduisant à l’effacement des opérations nécessaires à l’accomplissement de mesures ou de calculs. Les infrastructures informationnelles ont donc des conséquences éminemment politiques sur les définitions d’une activité qui sont embarquées dans leur constitution matérielle, certaines dimensions de cette activité étant jugés centrales et valorisées, quand d’autres sont considérées comme marginales et ignorées.

Ainsi, l’une des opérations bibliométriques les plus courantes consiste à calculer le nombre de citations d’un texte. Or elle a supposé une véritable standardisation du format des textes scientifiques puisqu’elle a impliqué l’homogénéisation préalable des formes de référencement et l’isolement des références (Bazerman 1988). Elle a également mis en

---

[2] “Une politique du formalisme s’appuie sur plusieurs choses. Tout d’abord, une représentation formelle est une abstraction : elle dégage des propriétés à partir d’une situation particulière. Ensuite, c’est une simplification : elle réduit la complexité des situations réelles pour les rendre formellement (généralement, mais pas exclusivement, mathématiquement) malléable. Enfin, et le plus important, toute représentation formelle contient des choix à propos de ce qu’il faut conserver (ce qui est important) et ce qu’il faut jeter. Tous ces choix sont politiques”.

équivalence des formes d'intertextualité très différentes : critique mesurée ou radicale d'une pensée ou d'un résultat, appui sur une méthode standardisée décrite ailleurs, usage d'autres textes comme des données pour une revue de littérature ou une méta-étude, simple recopie de sources référencées ailleurs (Brooks 1986 ; Liu 1993). Enfin, elle a demandé de brouiller la différence entre référence et citation : l'acte de référence relève d'un auteur donné alors que la citation est une nouvelle propriété, éventuellement calculable, du texte source. Selon P. Wouters (1999), ce renversement a radicalement modifié les pratiques de référencement et littéralement créé une nouvelle "culture de la citation". C'est seulement après ces différentes opérations que la mise en équivalence entre l'importance d'un texte et son niveau de citation est rendue possible, et qu'une évaluation à distance peut s'exercer, sans nécessiter une connaissance précise du domaine ou même la lecture du texte en question.

Partant de cette conception matérielle et distribuée, nous analyserons ici comment les technologies de calcul des productions scientifiques articulent plusieurs éléments, très souvent amalgamés dans les discours et les pratiques quotidiennes. Ces technologies combinent notamment trois composants des infrastructures informationnelles qui soutiennent l'évaluation quantitative des productions scientifiques. Nous prendrons soin de les distinguer pour mieux en saisir tant les conditions d'émergence et les singularités que la portée pragmatique et politique.

Tout d'abord, les algorithmes sont des formules mathématiques : de simples dénombrements, des ratios entre deux quantités ou, parfois, des calculs algébriques plus complexes. Les nombreux algorithmes, développés à partir des années 1940 et qui prolifèrent depuis quelques années (Bollen et al. 2009), portent sur les articles, les revues, les auteurs et les institutions de recherche, dont ils cherchent à mesurer la production ou les usages. Ils sont généralement proposés par des chercheurs en bibliométrie ou en scientométrie, mais peuvent également émaner d'autres chercheurs ou institutions (universités, organismes de recherche, ministères…). À titre d'exemples récents, on peut évoquer la définition d'un "*successful paper*" comme "un article qui a reçu plus de citations que le nombre de références bibliographiques qu'il utilise" (Kosmulski 2011) ; ou celle du "taux de collaborations internationales d'une revue", calculé à partir de la nationalité des institutions des auteurs d'articles parus dans cette revue. De manière plus ancienne, le "taux d'acceptation" d'une revue, mettant en rapport le nombre d'articles publiés avec le nombre d'articles soumis, a rencontré un certain succès, notamment en gestion, de même que le "taux d'articles non cités" qui mesure la partie "morte" de la production d'une revue en divisant ces derniers par le nombre total d'articles publiés sur une période donnée.

Ensuite, les jeux de données constituent des sources organisées de documents (bases de résumés, d'articles, de revues, d'ouvrages, d'auteurs…) mis en séries et classés selon différents critères. Selon leur structure, ces jeux de données peuvent se voir appliquer différents algorithmes. Ils peuvent être construits expressément, à titre individuel ou collectif, pour les besoins d'une recherche, mais, comme d'autres applications dans divers secteurs économiques, ils font aussi l'objet d'investissements lourds de la part d'organismes publics ou d'entreprises privées. Par exemple, l'Institut National de la Recherche Agronomique (INRA) s'est doté en 2006 d'une archive ouverte comprenant l'ensemble des productions de ses chercheurs et a rendu ce jeu de données accessible et interrogeable par des protocoles standards, sous la forme de la base PRODINRA. Sur un temps plus long, l'Institute for Scientific Information (ISI), sur lequel nous reviendrons en détail, est une entreprise privée fondée en 1959 qui produit encore aujourd'hui de nombreux jeux de données et développe des services payants.

Enfin, les outils bibliométriques associent de manière relativement pérenne certains jeux de données avec des algorithmes particuliers. Ces outils reposent sur des combinaisons plus ou moins complexes entre formules et données, et leur degré d'ouverture est variable (accès aux données du calcul, accès aux résultats, protection intellectuelle des outils). Ainsi, les taux de collaborations internationales des revues sont calculés de manière routinière sur le

jeu de données Scopus, une infrastructure privée appartenant à Elsevier, mais ces taux sont rendus publics par le groupe de recherche SCImago. À l'inverse, le taux d'acceptation d'une revue est fondé sur l'exploitation de données que chaque revue possède seule, mais qu'elle peut mettre en avant et diffuser pour fonder son excellence scientifique (Macdonald et Kam 2007).

Cette distinction s'avère pertinente pour saisir le succès de certains algorithmes, jeux de données, et outils bibliométriques en tant que composants clés des infrastructures informationnelles dans les pratiques d'évaluation de la recherche scientifique. Elle est également utile pour repérer différentes configurations dans lesquelles ces éléments s'articulent de façon spécifique. Nous mettrons en évidence ici trois configurations particulières. Une première émerge à partir du milieu des années 1960, prend appui sur les développements de la scientométrie, et atteint son apogée, au moins en sciences biomédicales, au cours des années 1990. Elle est caractérisée par la position monopolistique d'un unique producteur de données, l'ISI, proposant une palette d'outils bibliométriques fondée sur des algorithmes principalement développés en interne. Parmi ces outils, le *Journal of Citation Reports* (JCR) a progressivement acquis une position centrale dans le paysage de l'évaluation scientifique.

Une deuxième configuration se dessine au milieu des années 2000 avec l'entrée en scène d'autres producteurs de données, le succès d'algorithmes inédits et le développement découplé d'outils bibliométriques. Elle se caractérise par une multiplication des instances proposant des instruments de mesure et un important renouvellement de l'offre des dispositifs destinés à évaluer les entités du monde scientifique.

Au cœur de ce mouvement, la place prise par des outils de mesure alternatifs jette potentiellement les bases d'une troisième configuration qui s'esquisserait sous nos yeux. Celle-ci semble marquée par une triple rupture : l'étalon de mesure historique que sont les revues est concurrencé par un déplacement des mesures vers les articles, l'importance des usages se substitue à celle accordée aux productions, et le cercle des pairs s'élargit à d'autres catégories d'acteurs. Autrement dit, quand la première configuration est confinée au périmètre d'une science des mesures, la deuxième, puis la troisième, ouvrent considérablement l'horizon toujours plus vaste des métriques de la recherche.

De l'analyse des citations au *Journal of Citation Reports*

En 2009, le rédacteur en chef du *Journal of Sleep Research* commençait ainsi un éditorial, intitulé "*the race for the impact factor*", où il décrivait l'attente générée par la parution annuelle du JCR :

> "Chaque année, pendant la deuxième quinzaine du mois de juin, les rédacteurs en chef et les éditeurs de revues deviennent nerveux et agités, et même peut-être insomniaques, attendant impatiemment la publication du facteur d'impact de leur revue (JIF) pour l'année en cours par Thomson Reuters (auparavant Thomson Scientific). Ces dernières années, le JIF a acquis une influence légendaire sur le prestige des revues et les articles qu'elles publient, qui a engendré une "course pour le facteur d'impact". Les rédacteurs en chef essaient d'augmenter leur JIF autant que possible et les auteurs s'efforcent de publier leurs articles dans les revues ayant le plus grand JIF possible" (Lavie 2009).

Dans l'ensemble des sciences biomédicales et dans une partie d'autres domaines, sciences physiques et sciences sociales, les commentaires sur les résultats du JCR sont devenus monnaie courante. Aujourd'hui vendu comme un moyen "objectif et systématique d'évaluer de manière critique les revues les plus importantes au niveau mondial, à l'aide d'informations

quantifiables fondées sur des données de citation"[3], le JCR est pourtant le fruit d'une trajectoire faite d'incertitudes. Dans cette section, nous allons revenir sur la genèse de cette pratique dominante de l'évaluation quantitative des revues scientifiques, en distinguant successivement les ingrédients agrégés dans l'algorithme du facteur d'impact, les réalisations graduelles de l'ISI comme producteur de données, et la lente diffusion de l'outil bibliométrique JCR.

Aux sources de l'algorithme : montrer l'influence par les citations

Contrairement à la présentation qui en est régulièrement donnée de nos jours, le facteur d'impact des revues scientifiques ne résulte pas de l'élaboration cohérente et linéaire d'un instrument d'emblée destiné à l'évaluation de la recherche. Comme l'ont montré Archambault et Lariviere (2009), ses conditions de production ont infléchi non seulement sa forme mais aussi son essor ultérieur. Il a été conçu dans une tradition des sciences de la documentation, dont le public était les bibliothécaires des universitaires américaines. Dans ce cadre, l'influence d'une revue se mesure à l'aulne de ses usages par un ensemble de revues, sous la forme de comptages ultérieurs de ses citations. C'est donc la quantité de citations qui marque l'influence.

Dans leur étude pionnière, Gross et Gross (1927) ont élaboré une méthode fondée sur le nombre total de citations, afin de disposer d'un moyen objectif de sélection des revues à acheter et à conserver, pour aider à gérer les collections de bibliothèques. Face à l'accroissement et la dispersion des productions, cette méthode permettait notamment de repérer un petit nombre de "revues centrales" (*core journals*) dans divers segments de la littérature scientifique (Bradford 1934). Cependant, le nombre total de citations d'une revue est fortement corrélé au nombre total d'articles. Rompant avec la tradition fondée par Gross et Gross, Raisig (1960) a proposé d'éliminer cet effet de structure en adoptant un ratio : le nombre d'articles cités, qu'ils le soient une fois ou cent fois, divisé par le nombre total d'articles publiés. Ce ratio définit son indice de Research Potential Realized (RPR), atteignant au maximum 1 pour une revue dont tous les articles seraient cités. C'est dans cette ligne que se situent Garfield et Sher (1963) quand ils calculent pour la première fois ce qu'ils appellent le "*journal impact factor*" (JIF), mais ils attribuent au numérateur le nombre total de citations d'une revue, et non le nombre d'articles cités. Cette modification a des conséquences importantes : du point de vue du RPR, une bonne revue est une somme d'articles dont chacun est ultérieurement cité, s'inscrivant pleinement dans la connaissance scientifique partagée ; dans la perspective du JIF, l'ensemble de la revue constitue l'entité dont on mesure l'impact, quels que soient les écarts entre les différents articles publiés.

Encore faut-il déterminer la période à prendre en considération pour recueillir les citations. En 1968, Martyn et Gilchrist ont développé une analyse montrant que 26% des citations contenues dans les articles étudiés en 1965 portaient sur des articles parus en 1963 et 1964 (Archambault et Larivière 2009). C'est en prenant appui sur ce résultat que Garfield (1972) a systématiquement calculé le JIF de 1969 pour 152 revues à partir des citations des articles parus en 1967 et 1968. Et il a avancé l'idée que le JIF est un représentant du comportement de l'article "moyen", renforçant la conception d'une revue comme unité et non comme somme d'articles. Mais les revues scientifiques généralistes ne sont pas composées que d'articles originaux de recherche : elles publient également des éditoriaux, communications techniques, lettres, notes, compte-rendus d'ouvrages... Comme le souligne Bensman (2007), Garfield a également comptabilisé ces types de publication dans le dénominateur du nombre total d'articles publiés. Il a même douté de la possibilité de construire une classification pertinente de ce matériel diversifié. Pourtant, quatre ans plus tard, lorsqu'il a publié pour la première fois le recueil systématique des JIF dans le *Journal of Citation Reports* (JCR), il

[3] http://thomsonreuters.com/products_services/science/science_products/a-z/isi_web_of_knowledge/ (consultée le 06.06.2012).

indique que le dénominateur de son algorithme est la somme des articles “citables” publiés par une revue, sans en donner la méthode de comptage ou de sélection (Garfield 1976).

Faire des données bibliographiques une entreprise rentable : la genèse de l’*Institute for Scientific Information*

Le JCR n’est pas un document isolé, il fait partie d’une ligne de produits développés par l’ISI, l’entreprise privée fondée par E. Garfield en 1959. En effet, celui-ci n’est pas seulement un linguiste bibliomètre élaborant des indices de mesure ; c’est également un entrepreneur, consultant en documentation intéressé par l’automatisation, qui a progressivement bâti une industrie dédiée aux analyses quantitatives de l’information scientifique. Certains de ces jeux de données ont même rapidement acquis de l’importance, à l’instar des *Current Contents* qui consistaient en une édition papier hebdomadaire des sommaires de plusieurs centaines de revues quelques semaines avant leur parution. Outre la compilation des sommaires, les *Current Contents* comprenaient un index des sujets thématiques et un index des auteurs avec leur adresse, de façon à ce que les lecteurs puissent leur demander des tirés-à-part des articles. Et c’est en partie à l’aide des profits dégagés par cette publication payante que Garfield a pu financer, non sans difficultés, la production d’un jeu de données à l’histoire particulièrement tortueuse : le *Science Citation Index* (SCI).

Wouters (1999) a bien montré que l’ambition première de Garfield était de réaliser un index de citations sur le modèle du *Shepard’s Citator*, une base de données couvrant l’ensemble des décisions judiciaires américaines et permettant de savoir, dans le système jurisprudentiel américain, si une règle de droit est encore valable. Mais la quantité d’articles scientifiques et leur dispersion étant bien plus vastes que celles des décisions judiciaires, le SCI n’a pu voir le jour qu’avec l’appui motivé de plusieurs personnes relais auprès d’institutions scientifiques majeures telles que les National Institutes of Health et la National Science Foundation (NSF)[4]. Le soutien des généticiens G. Allen et J. Lederberg a notamment permis à Garfield de bénéficier de plusieurs financements et de placer son projet au centre d’intenses débats sur la politique de la recherche et l’information scientifique au début des années 1960 aux États-Unis.

Le SCI a vu le jour en 1964 dans une version minimale. Son coût demeurait tel que, dans un premier temps, il a manqué de mettre l’ISI en faillite, les agences fédérales ayant cessé leur financement. Imaginé comme un outil sophistiqué d’exploration de la littérature scientifique, le SCI a connu le succès pour un usage imprévu : en tant qu’infrastructure informationnelle, il rencontrait les préoccupations d’une “science de la science” prônée par Solla Price (1963) et équipait les travaux de sociologie des sciences (Cole et Cole 1971). Mais il accompagnait également la naissance d’une spécialité à la croisée des sciences sociales et de l’expertise bureaucratique, la scientométrie, engendrant des représentations inédites de la science qui ont influé sur la politique de la recherche. En revanche, le SCI a reçu un accueil très mitigé pour son usage prescrit.

Ce succès relatif du SCI tient dans sa grande originalité à l’époque : celle de couvrir un vaste ensemble de domaines scientifiques. L’automatisation, rendue possible par les ordinateurs, soutenait cette volonté de décloisonnement. En renonçant aux découpages disciplinaires, ce jeu de données permettait de signaler des références pertinentes dans des domaines inattendus aux chercheurs qui connaissaient bien le leur. En revanche, le SCI n’était pas voué à accueillir l’ensemble des revues. C’est ici que l’algorithme du facteur d’impact prenait tout son sens : après avoir identifié les “revues centrales” dans chaque domaine à l’aide du nombre total de citations, Garfield s’est servi du facteur d’impact pour sélectionner les revues complémentaires destinées à figurer dans le SCI (Bensman 2007), et notamment celles spécialisées dans les articles de synthèse.

---

[4] Nous résumons ici une histoire retracée par Wouters (1999), dont seule une synthèse très partielle est disponible en français (Wouters 2006).

Dans une logique d'extension, et poussé par l'intérêt que le SCI suscitait chez les scientomètres et les sociologues des sciences, l'ISI créait en 1972 un nouveau jeu de données, le Social Sciences Citation Index (SSCI). La liste des revues de sciences sociales incluses résulte d'une sélection à partir de trois sources : celles fortement citées dans le SCI, celles distinguées par la littérature en sciences de l'information, et celles désignées par des chercheurs sollicités par l'ISI (Garfield 1972). Poursuivant son processus d'extension, l'ISI envisageait de produire un jeu de données spécifique aux arts et humanités. Les pratiques de citation dans certaines disciplines sont toutefois retorses : on fait référence de manière implicite, ou encore les sources premières sont autant, voire plus importantes, que la littérature mobilisée (Garfield 1977). Mais, si la publication du *Arts & Humanities Citation Index* en 1978 a impliqué une modification de la nature des citations considérées, c'est bien le modèle du SCI qui achevait son extension.

Les jeux de données produits par l'ISI couvraient ainsi l'ensemble des disciplines universitaires à partir d'une sélection raisonnée de revues supposément centrales, et conçues comme les entités privilégiées du processus d'étalonnage.

Un outil bibliométrique à succès : le *Journal of Citation Reports*

L'apparition du JCR en 1975 a marqué un tournant. C'est en effet un outil bibliométrique qui regroupe, sur les jeux de données produits par l'ISI, les résultats de plusieurs algorithmes (nombre total de citations, "*immediacy index*" et JIF) portant directement sur les revues, et non sur les articles, les auteurs ou les objets de recherche. Comme le rappelle l'exemple donné précédemment, le JCR paraît annuellement et délivre des mesures régulières sur les revues qui y sont intégrées.

Cette nouvelle offre a donné lieu à des usages imprévus, à la suite notamment de l'intérêt suscité par les données calculées pour le JCR aux yeux d'autres acteurs. Tout d'abord, des revues qui n'étaient pas jusque-là considérées comme centrales s'en sont emparées pour souligner leur grande influence : dès sa première publication, des éditoriaux d'auto-satisfaction fleurirent dans les revues spécialisées dans la publication des articles de synthèse.

> "Une analyse récente […] de 50 revues botaniques mondiales a montré que, parmi les articles de botanique cités, l'*Annual Review of Phytopathology*… se classait deuxième…, en dépit… d'un taux d'auto-citation très faible. Cela en dit beaucoup sur la nature et la portée des articles de synthèse" (*Annual Review of Phytopathology* 1975).

Ensuite, des chercheurs ont également utilisé les résultats du JCR pour tenter de dessiner de nouvelles hiérarchies internes à leur discipline. Par rapport à des classifications construites antérieurement sur les jugements agrégés des pairs, démarche qui s'était répandue au début des années 1970 (Pontille et Torny 2010), celles élaborées à partir du JIF recueillaient de nombreux soutiens. Le JCR s'installait progressivement comme une entité naturelle pour l'ensemble des acteurs du monde académique, dont l'immense majorité était auparavant peu férue de bibliométrie. Par exemple, des universitaires italiens ont dénoncé l'injustice du processus de recrutement des professeurs dans leur pays en prenant appui sur les lieux de publication des différents candidats et leur "valeur" dans le JCR (Fabbri 1987 ; Aiuti et al. 1991). La nouvelle centralité du JCR est également démontrée par la tentative d'une coalition d'acteurs institutionnels d'y introduire plus de revues publiées en Afrique, en Asie et en Amérique du Sud afin d'obtenir une meilleure visibilité pour les chercheurs de ces pays (Moravcsik 1985). Une telle intégration devenait d'autant plus cruciale que différentes institutions se sont progressivement approprié cet outil bibliométrique en tant qu'indice de la qualité des publications des équipes de recherche, sans tenir compte du caractère très américano-centré du JCR (Pons-Novell et Tirado-Fabregat 2010).

Enfin, la réussite du JCR est devenue visible à travers les critiques récurrentes de son pouvoir, dont une première version élaborée apparaît en 1993 sous la plume d'un éditeur de Science prenant notamment l'exemple d'un directeur de laboratoire québécois.

> "[Il] répartit les ressources et promotions dans le laboratoire sur la base d'un système de notation dans lequel la taille des bourses du chercheur compte pour 40%, les performances des doctorants et des post-docs pour 20%, et les citations pour les 40% restants. Ce système, affirme Labrie, incite les chercheurs à publier dans les meilleures revues, qui leur donneront le plus grand nombre de citations. Quand on l'interrogeait pour savoir s'il considérait ce système de notation quantifiée quelque peu impersonnel et froid, Labrie répondit qu'il n'était 'pas plus cruel que la vie elle-même'" (Taubes 1993).

Alors que les scientomètres sont toujours demeurés critiques à propos de tels usages d'outils bibliométriques (Van Raan 2005 ; Gingras 2008), certains chercheurs ont eux-mêmes valorisé le JIF. Ce constat est régulièrement renouvelé dans la décennie suivante : de nombreux pays ont adopté des mesures d'incitation ou d'évaluation fondées sur le JIF, et des chercheurs modifient leurs pratiques de publication et de citation afin d'être bien positionnés dans les hiérarchies inédites favorisées par ces technologies d'évaluation.

Ce mouvement d'appropriation s'est accéléré avec la mise en ligne des services proposés par l'ISI, et notamment l'élaboration du *Web of Science* (WoS), un dispositif permettant d'interroger les différents jeux de données et le JCR à partir de n'importe quelle bibliothèque universitaire possédant un abonnement. Au début des années 2000, l'ISI était donc en situation de quasi-monopole pour la production des données à destination d'outils bibliométriques[5]. Parmi la large gamme des produits que l'entreprise propose, ce sont ses usages multiples qui ont favorisé la quasi exclusivité du JCR, et de son algorithme JIF, comme mesure de la qualité scientifique des revues et, par dérivation, des institutions et des chercheurs. Même si de nombreuses revues regrettaient cette situation et pointaient les limites de cet outil bibliométrique (Opthof 1997 ; Seglen 1997), elles constataient simultanément son caractère incontournable.

Toutefois, cette position dominante a rapidement évolué au cours des années 2000 avec l'émergence d'autres producteurs de données, la percée d'algorithmes inédits et la diffusion d'outils bibliométriques alternatifs. Cette multiplication des entités se positionnant sur la scène de l'évaluation scientifique a déclenché le passage d'un monde monopolistique à un monde pluraliste ouvert sur les métriques de la recherche. Autrement dit, elle a engendré une toute nouvelle configuration.

L'entrée en scène d'autres producteurs

Le monopole acquis par l'ISI résulte d'un lent travail d'accumulation monétaire, logistique et commerciale. Garfield a promu et défendu de manière continue ses outils auprès de financeurs, de bibliothécaires, d'universités et de revues. Pourtant cette position dominante n'était pas nécessairement assumée par Garfield lui-même, qui appelait à l'utilisation de tout jeu de données à des fins de production de connaissances (Garfield 1979). D'autres entrepreneurs ont tenté de faire exister leurs propres algorithmes, jeux de données et outils bibliométriques. C'est notamment le cas de F. Narin qui a bâti en 1968 Computer Horizons Inc. (CHI), dont l'un des objectifs était la production d'une bibliométrie à visée explicitement évaluative qui, à ce titre, a notamment reçu des financements de la NSF (Narin 1976). Nous y reviendrons. Il a cependant fallu attendre le début des années 2000 pour que deux entreprises apparaissent comme de véritables concurrents à l'ISI : l'éditeur scientifique néerlandais Elsevier, et le géant du web américain Google.

---

[5] Ce succès a fait de l'ISI une entreprise rentable qui a été racheté par le groupe Thomson Scientific & Healthcare en 1992.

Scopus d'Elsevier : de l'édition scientifique au jeu de données bibliographiques

Elsevier est un éditeur scientifique, fondé en 1880 à Rotterdam, qui a acquis à la fin du XX^{e} siècle, par le biais de rachats successifs, une position dominante, mais non monopolistique, dans le domaine biomédical. En 2001, l'entreprise a créé Scirus, un moteur de recherche spécialisé dans les contenus scientifiques qui permettait, via un accès web, de faire des requêtes en texte intégral sur l'ensemble des articles publiés non seulement par Elsevier, mais aussi par d'autres éditeurs commerciaux. Pour ce faire, l'entreprise a utilisé des bases de données qu'elle possédait (EMBASE, GEOBASE, COMPINDEX), ainsi que d'autres construites par ailleurs (PubMed). Même sans un abonnement payant à ses bases de résumés d'articles, tout chercheur pouvait donc utiliser Scirus à des fins d'exploration bibliographique. La devise de ce moteur de recherche, "*for scientific information only*", réincarnait l'objectif initial de l'ISI, mais dans une acception différente. Alors que pour Garfield, il s'agissait de construire de novo des jeux de données permettant de repérer l'information pertinente dans l'océan des publications disponibles, pour Elsevier, il suffisait de sélectionner, parmi les jeux de données existants[6], ceux contenant des informations scientifiques pour les agréger.

Elsevier a ensuite entrepris la construction d'un second instrument en relation avec des universités et des chercheurs en bibliométrie. Le 3 novembre 2004, l'entreprise a présenté Scopus, un service payant qui, outre l'intégration des fonctions de Scirus, donnait accès au texte intégral ainsi qu'aux articles citants et cités. Ce nouveau produit constituait la première véritable concurrence pour l'ISI WoS, en raison de son caractère généraliste : Elsevier et des commentateurs soulignèrent les points communs et les différences entre les deux jeux de données, par exemple l'indexation de l'ensemble des auteurs d'un même article et non du premier uniquement, ou la présence plus importante de productions dans d'autres langues que l'anglais (Jacso 2004). Ce modèle semi-ouvert, s'appuyant sur des échanges continus avec les utilisateurs, se retrouve dans les développements successifs de Scopus, comme l'ajout, par les utilisateurs, de liens hypertextes vers le texte intégral de leurs publications. Ce sont également les utilisateurs qui peuvent suggérer, via une page web spécifique[7], des titres de supports à inclure dans Scopus, ces derniers étant sélectionnés par un comité, composé de 20 scientifiques et 10 bibliothécaires, qui n'a pas adopté une logique de quota contrairement à l'ISI[8].

Alors que les jeux de données de l'ISI se fondaient sur un double argument – l'excellence de leur contenu et la réduction de la masse d'informations – Scopus est un dispositif à vocation d'exhaustivité. Loin de refuser des revues au contenu "exotique" comme l'ISI a pu le faire dans les années 1980, Elsevier ne cesse au contraire de mettre en avant la taille de ses jeux de données, son extension vers des domaines peu couverts jusqu'alors (e.g. sciences humaines et sociales, brevets, supports dans d'autres langues que l'anglais). Il ne s'agit pas d'un enjeu simplement cognitif, mais également commercial puisqu'Elsevier démarche les universités ou les pouvoirs publics pour leur vendre ses jeux de données en affirmant que Scopus offre une meilleure couverture de l'activité scientifique et de son rayonnement et, par là même, permet une meilleure évaluation de cette activité.

---

[6] Il faut préciser que, dans ce cadre, les sites web sont traités comme des données et que Scirus assurait éliminer ceux dont les contenus n'étaient scientifiques.

[7] http://suggestor.step.scopus.com/suggestTitleAgreement.cfm (consultée le 30.06.2012).

[8] Les conditions d'inclusion comprennent l'existence de résumés en anglais, l'évaluation par les pairs et la parution régulière. L'ISI, puis Thomson, ont régulièrement indiqué ne retenir que 8 à 12% des demandes provenant de revues, alors qu'Elsevier focalise l'attention sur le très grand nombre des revues sélectionnées par son comité.

Google Scholar : du moteur de recherche au moteur pour la recherche

En novembre 2004, l'entreprise Google a mis en place un jeu de données associé à un moteur de recherche spécialisé, Google Scholar. Tout comme Scirus, il est accessible à tous via un site internet et permet de faire des requêtes sur les auteurs, les titres, les résumés. Une différence concerne cependant le critère "*for scientific information only*" qui n'était pas repris pour circonscrire la recherche de l'information pertinente. Google Scholar ne sépare pas a priori les sites "académiques" des ceux qui ne le sont pas, mais utilise l'intégralité des jeux de données sur lesquels s'appuie le moteur généraliste, y repérant automatiquement des motifs lexicaux de références bibliographiques.

Ce choix de l'extension maximale est confirmé dans les développements ultérieurs : le contenu des ouvrages numérisés dans Google Books était accessible dans le jeu de données Scholar fin 2004, et Google a entrepris un travail de numérisation de fonds de revues scientifiques en 2007, en complément d'accords avec des éditeurs scientifiques commerciaux. Google a également proposé sur son site des procédures permettant aux bibliothécaires et aux dépôts d'archives d'inclure leurs fonds dans les données Scholar. Dans la droite ligne du modèle économique de Google, combinant gratuité des usages et vente d'espaces publicitaires plus ou moins personnalisés, tout texte numérisé est potentiellement indexé dans le jeu de données Scholar. Les utilisateurs ont alors accès à des productions en texte intégral, archivées sur des sites personnels ou dans des dépôts institutionnels, quel que soit leur format informatique, leur langue ou leur genre (rapport, papier de travail, acte de colloque, chapitre d'ouvrage, tribune dans la presse...), que leur contenu soit ou non réservé à des abonnés. De plus, le jeu de données Scholar est indexé de telle sorte qu'il permet le suivi des productions citantes et citées (Giles 2005), ce qui en fait un concurrent de facto à l'ISI WoS et au tout nouveau Scopus, même en l'absence de commercialisation directe d'un "produit" Google Scholar.

Cet espace concurrentiel est très visible dans les travaux de bibliométrie et de scientométrie : à partir de 2005, des articles mettent en scène les trois jeux de données, avec différents objectifs. Certains cherchent à vérifier leur "qualité" : couvrent-ils effectivement plus de sources que celui de l'ISI ? Comment traitent-ils les doublons ? Franchissent-ils la barrière du "non-scientifique" ? Proposent-ils plus de fonctions ? (Gardner et Eng 2005 ; Burnham 2006 ; Moya-Anegón et al. 2007). D'autres opèrent une comparaison directe entre les trois jeux de données, sur une discipline ou un domaine de recherche donné, pour en montrer la cohérence ou souligner les écarts générés en termes bibliométriques (Meho et Yang 2006 ; Norris et Oppenheim 2007). Cette comparaison peut également déboucher sur la construction de jeux de données combinés, ou sur leur hiérarchisation en fonction des usages (Falagas et al. 2008 ; Bar-Ilan 2010).

L'activité de producteurs de jeux de données à destination des chercheurs s'est fortement diversifiée ces dix dernières années, Elsevier et Google n'étant que les deux acteurs les plus visibles et pérennes. On aurait pu aussi évoquer le cas de Microsoft, le producteur américain de logiciels, qui a proposé avec Live Search Academic un service proche de ceux de Google Scholar entre 2006 et 2008. Cette pluralité, inséparable du développement de l'internet à haut débit, a trois caractéristiques : l'extension de la liste des productions scientifiques visibles, l'accessibilité des produits de la recherche, et le détachement grandissant entre support de publication et contenu, via notamment des usages de "*citation tracking*". Mais l'accessibilité a également un effet en retour sur les algorithmes : puisque des données sont désormais plus facilement disponibles, des algorithmes anciens ou nouveaux peuvent être aisément testés ou mis en œuvre à une large échelle.

Le succès de “nouveaux” algorithmes

L’extension progressive de la scientométrie et du nombre de revues spécialisées (Scientometrics, JASIST, Research Evaluation, Journal of Informetrics,...) se traduit par la proposition continue de nouveaux algorithmes. Ceux-ci peuvent être testés sur de petits jeux de données, sur quelques disciplines ou sur une gamme de revues. Leur carrière s’infléchit considérablement lorsqu’ils deviennent la cible de critiques récurrentes, telle la prise en compte du coautorat, des auto-citations ou du nombre de références par article, et qu’ils font l’objet de modifications progressives. Parmi toutes ces propositions, un petit nombre connaît un succès important par leur association à divers jeux de données et, finalement, leur incorporation dans des outils bibliométriques standardisés. Nous focaliserons l’attention ici sur deux algorithmes aux caractéristiques contrastées : le premier a une longue trajectoire interne à la scientométrie, le second a connu un succès fulgurant après avoir été proposé par un physicien.

De l’*influence weight* au *PageRank* bibliométrique : une citation ne vaut pas une citation

Nous avons déjà évoqué la tentative de F. Narin et son entreprise CHI de concurrencer l’ISI de Garfield au milieu des années 1970. Il défendait un modèle de mesure de l’influence radicalement différent de celui du JIF : toutes les citations ne se valent pas. Avec un tel principe, Narin souhaitait tenir compte du prestige du lieu de citation. Contrairement à Garfield qui considérait équivalentes les citations réalisées au sein de sa sélection de revues centrales, il entendait marquer une nette différence entre elles, voire sur une gamme de revues plus large. Aussi, tout en utilisant les données du SCI, Narin a proposé des définitions de l’influence d’une revue qui lestent les citations d’un poids relatif à la revue citante : plus une revue est citée par une revue prestigieuse, plus elle est prestigieuse (Pinski et Narin 1976). Ce principe récursif de l’*influence weight* demande toutefois beaucoup de temps pour les calculs matriciels nécessaires à la production des résultats, et Narin les a publiés seulement pour les revues de physique et de biochimie présentes dans le SCI. Cette contrainte a limité les ambitions de CHI de couvrir rapidement l’ensemble des domaines scientifiques, contrairement à l’ISI. Narin a alors consacré son entreprise aux analyses bibliométriques des brevets (Narin 1994). Son algorithme est longtemps demeuré confiné à de petits groupes de scientomètres, seuls quelques chercheurs l’ayant utilisé dans des opérations d’évaluation[9].

Un quart de siècle plus tard, lorsque Page et al. (1999) proposent de mettre de l’ordre dans le world wide web en décrivant l’algorithme de classement de Google[10], *PageRank*, ils utilisent des principes mathématiques proches de ceux de l’*influence weight*. Ces similarités sont notées pour la première fois par Kleinberg (1999), lorsqu’il traite, dans un article désormais célèbre, de la quantification de l’autorité des sources. À partir de là, les discussions sur le *PageRank*, la classification des moteurs de recherche du web et ce que l’on commence à nommer les webometrics (Almind 1997) embarquent la portion d’histoire algorithmique de l’*influence weight*.

Dans un premier temps, ce succès a eu peu d’effets sur le domaine des études bibliométriques : les scientomètres soulignaient plutôt les différences entre les réseaux de pages du web et les intercitations dans les revues, jugeant avec réserve la possible réimportation de l’algorithme (Prime et al. 2002). Il faut attendre l’article séminal de Bollen et al. (2006), sobrement intitulé Journal Status, pour que l’*influence weight* soit au centre des

---

[9] En revanche, le principe d’hétérogénéité de la valeur des citations a été proposé dans certaines disciplines, de manière indépendante de l’*influence weight*. Pour l’économie, voir Liebowitz et Palmer (1984).

[10] Voir dans ce même numéro l’article de D. Cardon (2013).

débats. Prenant appui sur le succès de Google et du *PageRank*, les auteurs annonçaient une véritable révolution :

> "À l'avenir, le *PageRank*, et non le JIF de l'ISI, pourrait incarner notre représentation des revues et des articles. La transition d'un classement de revues fondé sur l'algorithme du JIF de l'ISI vers celui du *PageRank* indiquerait le passage d'un modèle d'évaluation fondé sur la popularité, i.e. la fréquence de citation, vers un modèle fondé sur le prestige, i.e. le prestige de ceux qui citent est pris en compte" (Bollen et al., 2006, p. 670-671).

Cette annonce de révolution a néanmoins fait place à une position plus réformiste puisque les auteurs proposaient d'adopter le Y-factor pour les revues, résultat de la multiplication du JIF de l'ISI par le *PageRank*. Ils reprenaient ainsi une conception algorithmique développée dans les années 1970 pour des classements de revues, où la valeur résulte du produit d'une mesure de popularité par une mesure de prestige (Pontille et Torny 2010). Après que l'article ait connu une grande visibilité suite à un compte-rendu dans Nature, de nombreux chercheurs se lancèrent dans le calcul "révolutionnaire" du *PageRank* pour les revues. Ce qui a produit deux effets : d'une part, des vagues de publications dans les disciplines les plus diverses, de la dermatologie à la biologie moléculaire et, d'autre part, le rapprochement entre communautés travaillant sur différentes formes de mesures (scientomètres, infomètres, webomètres) (Egghe 2005).

Mais l'algorithme a également été inclus dans deux outils bibliométriques stabilisés. Bergstrom (2007), un biologiste américain, l'applique sur les données ISI pour fabriquer l'Eigenfactor, qui mesure l'influence totale d'une revue, et l'ArticleInfluence, qui la normalise par article publié, reprenant ipso facto les propositions algorithmiques de Pinski et Narin (1976). Ces résultats sont disponibles, sans abonnement, sur le site www.eigenfactor.org. À partir de 2010, l'Eigenfactor entre dans le JCR, faisant donc désormais partie de la gamme des mesures sur les revues fournies par l'ISI, à côté du JIF. Parallèlement, Elsevier a passé un contrat avec une équipe de scientomètres espagnols, qui produit, sur les jeux de données de Scopus, le SCImago Journal Ranking (SJR) à partir de 2007, et est également disponible gratuitement sur www.scimagojr.com. Lorsqu'il est introduit dans le jeu de données Scopus, le SJR est présenté comme la mesure la plus pertinente de l'influence d'une revue. Selon les producteurs de données, l'*influence weight* joue donc soit le rôle d'un algorithme complémentaire, soit celui de substitution au JIF[11].

À la manière des jeux de données apparus dans les années 2000, ces nouveaux outils font l'objet de publications évaluant leur intérêt, leurs limites, les comparant systématiquement avec le JIF ou les présentant en parallèle avec ce dernier dans des classements de revues. Si, du point de vue algorithmique, il y a une véritable rupture avec la métrique du JIF, et si l'ouverture des jeux de données et la disponibilité des résultats les rend plus visibles, les outils SJR et Eigenfactor jouent un rôle similaire : ils fournissent une évaluation des revues académiques sous la forme d'un nombre et éventuellement d'un rang.

Le h-index : mesurer des auteurs à l'aide de citations

Le second algorithme à succès a opéré une rupture beaucoup plus radicale pour une triple raison : au lieu d'être issu du raffinement de mesures antérieures, il était totalement inédit ; il n'a pas été proposé par un scientomètre, mais par un physicien qui n'envisageait pas de le rapporter à d'autres algorithmes ; il n'agrégeait pas une mesure sur les articles en direction des revues, mais vers les auteurs. Certes, l'usage d'algorithmes portant sur des auteurs est ancien : Garfield prétendait régulièrement prédire les prix Nobel grâce au SCI et il avait ensuite développé le "*ISI highly cited*" présentant des résultats pour une toute petite fraction

[11] Les deux outils ont également pour différence la période considérée : l'Eigenfactor utilise une fenêtre de 5 ans pour les citations, le SJR de 3 ans.

de chercheurs[12] ; de plus, dans certaines disciplines comme l'économie et la gestion, existaient de longues traditions de classification des "*top authors*". Mais aucun algorithme n'avait encore été spécifiquement conçu pour évaluer uniquement les auteurs : le nombre total de citations, le nombre total d'articles, le nombre de citations par article étaient utilisés pour tout type d'entité (auteur, laboratoire, université, pays).

Le physicien J.E. Hirsch a proposé le h-index dans un article de travail déposé le 3 août 2005 sur arXiv, archive ouverte développée par des physiciens à Los Alamos[13]. Il y aborde le problème de l'évaluation globale d'un chercheur et souhaite subsumer l'ensemble de la carrière sous une mesure simple et pratique : un nombre entier. Pour cela, il considère qu'il faut prendre en compte sa production et les usages de cette production à travers une mesure de citation. Le nombre h est le plus grand nombre pour lequel h articles d'un auteur ont au moins h citations. Par exemple, pour 5 articles cités au moins 5 fois, le h est de 5 ; de même, pour 50 articles cités au moins 50 fois chacun, il est de 50. Cette proposition s'accompagne d'un modèle normatif du "bon chercheur" : comme le résultat de l'algorithme dépend beaucoup de la durée de la carrière, le ratio h divisé par le nombre d'années de carrière est un bon indicateur pour Hirsch.

> "Un h-index de 20 après 20 années d'activité scientifique caractérise un savant qui a réussi […] un h-index de 40 après 20 années d'activité scientifique caractérise les savants exceptionnels, qu'on ne trouve probablement que dans les meilleurs universités ou les grands laboratoires de recherche […] un h-index de 60 après 20 années, ou de 90 après 30 années, caractérise des individus réellement uniques" (Hirsch 2005, p. 16571).

Ne se situant pas dans la tradition scientométrique, l'algorithme de Hirsch opère différentes ruptures. Tout d'abord, il ne tient pas compte des revues dans lesquelles les articles sont publiés, en raison sans doute du fait que, pour les physiciens des hautes énergies, les revues servent davantage à archiver les connaissances qu'à annoncer des découvertes (Gunnarsdóttir 2005). Ensuite, il écarte les questions du nombre et de l'ordre des signataires, qui sont peu pertinentes en physique alors qu'elles sont cruciales en recherches biomédicales pour l'évaluation individuelle (Pontille 2004). Il mêle deux éléments considérés comme hétérogènes dans les traditions scientométriques : la production d'une part, l'usage d'autre part. Enfin, alors que les bibliomètres sont toujours très prudents en matière d'analyse individuelle et la réserve à des scientifiques "hors normes", Hirsch propose une mesure qui s'applique à tout chercheur, a minima en physique, et indique en particulier qu'elle serait utile pour l'allocation des fonds de recherche.

À l'inverse de l'algorithme de Narin dont le succès est très tardif, le h-index est marqué par une adoption fulgurante. Dès la mise en circulation de son manuscrit, Hirsch a reçu de nombreux commentaires de ses collègues physiciens sur son algorithme et, devant l'enthousiasme partagé, une archive ouverte spécialisée en physique des hautes énergies, SPIRES (Stanford Physics Information REtrieval System), implémentait l'algorithme sur son jeu de données deux semaines plus tard. Le lendemain, le 18 août 2005, un article de la revue Nature signalait la proposition de Hirsch et soulignait l'enthousiasme de ses collègues, tandis qu'un éditorial, intitulé "*Rating Games*", discutait la place respective des métriques et de l'évaluation par les pairs. Deux collègues de Hirsch assuraient une publication rapide de l'article dans la revue *Proceedings of the National Academy of Sciences*, en novembre 2005, achevant de montrer l'intérêt pour cette nouvelle mesure chez les physiciens (Hirsch 2005).

Cet algorithme a ensuite fait l'objet de discussions intenses entre scientomètres, mais contrairement aux débats sur le JIF ou d'autres produits de l'ISI, les débats ne sont pas rattachés à l'auteur de l'algorithme princeps. Certains insistent sur ses limites, ses variations d'un domaine à l'autre, l'avantage des disciplines où le nombre de signataires d'articles est

---

[12] Thomson-Reuters indique que moins de 0,5% des chercheurs étaient identifiés dans cet outil.

[13] Sur les propriétés et les enjeux de cette base de preprints, voir Gunnarsdóttir (2005).

important (Iglesias et Pecharromán 2007), ou proposent des “versions améliorées” comme le g-index (Egghe 2006) ou le v-index (Ferrara 2012). D’autres proposent une extension du h-index à divers objets, notamment les groupes de chercheurs et l’objet canonique de la bibliométrie, les revues, pour une période déterminée – h-5 pour une fenêtre de 5 ans par exemple (Braun et al. 2006). De nombreux travaux comparent les classifications des chercheurs en utilisant différentes métriques, dont le h-index, et une évaluation par les pairs (Van Raan 2006). D’autres enfin soulèvent la question de l’appariement de l’algorithme avec les différents jeux de données nouvellement disponibles, Scopus et Google Scholar (Bar-Ilan 2007).

Et c’est justement cet appariement qui est à l’origine d’un nouvel outil bibliométrique élaboré par Anne-Wil Harzing : Publish or Perish (PoP). Cette professeure de management à l’Université de Melbourne a mis en ligne en octobre 2006 un petit programme qui rend opérationnel le calcul du h-index. De la sorte, pour n’importe quel auteur dont le nom est entré par l’utilisateur, quelle que soit sa discipline, PoP calcule en un clic son h-index sur le jeu de données Google Scholar. Cet outil est téléchargeable gratuitement et a donc permis à l’algorithme de toucher des utilisateurs bien au-delà des publics spécialisés de la scientométrie. L’adoption de cet outil bibliométrique par les chercheurs et les institutions a été tellement rapide que le British Medical Journal a publié un article parodique décrivant les différentes pathologies qu’il génère (Horne et al. 2009). Entre-temps, Elsevier a inclus en mai 2007 le h-index dans Scopus ; de même Thomson Reuters a supprimé le “*ISI Highly Cited*” et intégré cet algorithme dans le WoS en 2008. Ainsi, en seulement trois ans, la mesure individuelle est devenue à la fois un sujet de recherche central dans la production d’algorithmes pour les bibliomètres, et une opération pratique s’appuyant sur des outils bibliométriques facilement accessibles à chaque chercheur.

L’apparition régulière d’algorithmes n’informe pas uniquement sur leurs conditions particulières de félicité. Elle renseigne également sur les principales caractéristiques de cette deuxième configuration. Celle-ci est tout d’abord marquée par la multiplication des acteurs qui produisent des technologies de calcul de la recherche scientifique. Depuis les années 1970, l’ISI présentait une offre diversifiée sous couvert d’un unique référent. Dorénavant les éléments des infrastructures informationnelles destinées à l’évaluation de la recherche émanent de différentes instances : plusieurs entreprises pourvoyeuses de jeux de données plus ou moins ouverts et accessibles, des chercheurs producteurs d’algorithmes au succès tardif ou parfois fulgurant, des producteurs d’outils bibliométriques présentés comme des combinatoires. À l’inverse de la position monopolistique tenue par l’ISI jusque dans les années 2000, le monde de “la” bibliométrie est désormais peuplé par plusieurs producteurs de données en concurrence et un oligopole d’algorithmes “stars” traversant l’ensemble des milieux scientifiques.

Cette deuxième configuration est ensuite caractérisée par un découplage notoire des métriques de la recherche, les technologies de calcul proliférant sans être nécessairement encapsulés les unes dans les autres. Des producteurs de jeux de données se spécialisent exclusivement dans cette activité, à l’instar de Scopus qui rend visible, sur son site, la contribution indépendante de *SCImago* en tant que concepteur de l’algorithme et producteurs des résultats[14]. Symétriquement, des chercheurs peuvent développer des outils bibliométriques sans produire d’algorithmes ni confectionner de jeux de données, comme nous venons de le voir avec le PoP d’A.-W. Harzing. Plus radicalement, un même algorithme donne des résultats différents parce que la teneur de son calcul même dépend du jeu de données sur lequel il s’applique. Si les chercheurs peuvent désormais créer leur page authentifiée sur Google Scholar pour donner à voir leur h-index, celui-ci sera vraisemblablement différent sur les données de l’ISI et de Scopus.

---

[14] Cette présentation a été reprise par Scopus lorsqu’Elsevier a introduit, en plus du SJR, l’algorithme SNIP développé par H. Moed à l’université de Leiden.

Enfin, cette deuxième configuration amorce une diversification des points d'application des technologies d'évaluation quantitative des productions scientifiques. Avec l'arrivée de nouveaux producteurs de données et l'élaboration d'algorithmes inédits, les outils bibliométriques élargissent la gamme des entités visées : non seulement la définition de ce qui constitue une revue prestigieuse s'est transformée avec la redécouverte de l'*influence weight*, mais les revues ne constituent plus le seul et unique étalon de mesure. Pour autant, les technologies de calcul qui dominent dans cette deuxième configuration demeurent focalisées sur un seul usage des productions scientifiques : on scrute uniquement les citations que reçoivent des publications dans d'autres travaux scientifiques (revues, ouvrages, rapports…). Autrement dit, la référence aux pairs, en tant qu'auteurs de textes, est omniprésente et exclusive.

L'émergence récente d'outils alternatifs préfigure une troisième configuration, marquée par une ouverture encore plus grande vis-à-vis des entités visées, des mesures mobilisées, et des types de participants pris en compte.

Vers des outils alternatifs

Comme on l'a vu dans les prolégomènes du développement du h-index, les archives ouvertes comme jeux de données et lieu d'implémentation des algorithmes ont constitué des espaces nouveaux de créations bibliométriques. Leur généralisation dans les années 2000 et la transformation des espaces de publication, et notamment le passage de revues en open access, ont été porteurs d'une diversification encore plus grande des technologies de calcul. En effet, plutôt que de simplement concurrencer le modèle de l'ISI comme producteur de données ou de promouvoir de nouveaux algorithmes portant sur les citations, ces acteurs ont opéré deux transformations majeures : d'une part, la production et le stockage décentralisé de jeux de données ; d'autre part, l'adoption de mesures alternatives des usages, et pas seulement des citations, portant exclusivement sur les articles.

RePEc : construire un producteur décentralisé de données disciplinaires

La force et l'originalité de l'ISI, nous l'avons vu, étaient de proposer des jeux de données multidisciplinaires. Parallèlement, il existait des initiatives de partage de références, de bases de résumés, voire de textes complets sur une base disciplinaire. On peut donner l'exemple des Chemical Abstract Services fondés par l'American Chemical Society en 1907, dont la collection de résumés est passée de 12 000 à 6 000 000 en un siècle et dont les services n'ont cessé de se diversifier par l'inclusion d'une nomenclature des substances, des textes complets des articles et de leurs citations, et d'une base de données de brevets. En sciences humaines et sociales, on peut citer PsycINFO, créé par l'American Psychological Association, qui a inclus en 2001 des outils portant sur les citations. Dans cette veine, le projet Research Papers in Economics (RePEc), entamé en 1996, est marquée par une particularité : c'est le premier qui adopte une architecture décentralisée et mondiale consacrée non pas aux articles publiés, mais d'abord aux *working papers*.

Conçu par un groupe d'économistes britanniques, RePEc s'est d'abord appuyé sur une première infrastructure de *working papers*, WoPEc, fondée en 1993 avant l'apparition du world wide web. Avec l'avènement de ce dernier, le projet articule deux éléments distincts. D'un côté, chaque institution, chaque laboratoire peut développer son propre serveur et rendre disponibles les textes qui y sont stockés. De l'autre, les productions archivées ne sont pas directement accessibles, mais interrogeables via des applications dédiées : outre WoPEc développé en Grande-Bretagne, citons IDEAS au Canada, RuPEc en Russie, qui ont accès à l'ensemble des serveurs configurés pour RePEc. Le jeu de données lui-même contenait uniquement des *working papers*, avant de s'ouvrir progressivement aux articles

publiés, des partenariats étant notamment conclus avec les principaux éditeurs scientifiques commerciaux[15]. Au 1er mai 2012, le jeu de données comprenait plus de 450 000 *working papers* et plus de 700 000 articles publiés, ainsi que 15 000 ouvrages référencés, 90% du matériel étant téléchargeable (gratuitement ou avec abonnement aux éditeurs concernés).

L'infrastructure informationnelle associe donc une décentralisation de la production et un stockage des données de base (articles, références, résumés...), allié à la construction d'un jeu de données unifié sur lequel de nombreux services sont bâtis (Karlsson et Krichel 1999). Parmi ceux-ci, plusieurs sont des outils bibliométriques. Le projet LogEC, hébergé par la Stockholm School of Economics, enregistrait des statistiques des usages de NetEc ou IDEAS et proposait dès 2001 un "*top download*" ou un "*top abstract views*" des textes inclus dans RePEc. Il mettait donc en œuvre des algorithmes issus des webometrics, dénombrant simplement l'accès à des pages web ou des fichiers. Pour l'une des premières fois, la mesure de l'usage par ces outils bibliométriques ne se faisait donc pas en analysant des productions scientifiques citantes, mais en comptant des visionneurs de résumés et des téléchargeurs de *working papers*.

Simultanément, un registre des pages personnelles, avec les coordonnées d'économistes, où ils stockaient leurs *working papers* a été développé. En 2003, il a pris un sens différent, puisqu'il s'agissait d'identifier les auteurs des documents de RePEc et de leur fournir des statistiques mensuelles sur les pages vues et les téléchargements. Enfin, en 2004, à partir d'algorithmes développés par des informaticiens sur l'archive ouverte CiteSeer, les promoteurs de RePEc ont construit une mesure des intercitations sur l'ensemble du jeu de données. Cela permet d'appliquer des algorithmes déjà largement partagés comme le JIF ou l'*influence weight* pour les revues, pour les collections d'ouvrages ou pour les collections de *working papers*, les calculs s'effectuant bien évidemment sur la base des références des documents inclus dans RePEc. Il est également possible d'obtenir le classement des "*top authors*" en appliquant 35 algorithmes distincts sur le jeu de données, allant du nombre de pages publiées, qui tient compte du nombre de coauteurs, au nombre total de citations, en passant par le h-index, le "*top download*" ou le "*top abstract views*"[16]. De même, des agrégations sur les 1 779 institutions répertoriées sont disponibles. L'ensemble des divers classements est mis à jour et archivé une fois par mois.

Les outils bibliométriques de RePEc articulent donc un jeu de données très ouvert à une multiplicité d'algorithmes. Ces derniers peuvent provenir des institutions pionnières dans les deux configurations décrites plus haut ou être relativement spécifiques, comme le nombre d'auteurs RePEc citant un économiste donné. Mais le système RePEc ne hiérarchise pas les critères, sa présentation et son ergonomie assume un pluralisme des outils, chacun étant également disponible. Au traitement indifférencié des mesures correspond une absence de présélection des données : à la manière de Google Scholar, tout document déposé sur un serveur alimente le jeu de données, même s'il n'est jamais publié dans une revue ou téléchargé par les utilisateurs de RePEc. De plus, IDEAS ou RuPEc étant totalement libres d'accès, les visionneurs ou téléchargeurs, tout en étant décomptés, peuvent être des usagers non académiques.

Altmetrics : proposer un panier de mesures sur les articles

À la manière de RePEc qui propose une multitude d'options, le développement des réseaux sociaux tels Twitter et FaceBook s'est accompagné d'un engouement pour les mesures de l'influence, de l'impact ou du prestige. Au-delà des pages web professionnelles et des blogs, des réseaux sociaux spécifiquement dédiés aux chercheurs académiques sont désormais disponibles (e.g. Academia.edu, Mendeley.com, Peerevaluation.org) qui permettent eux aussi d'effectuer divers comptages et d'en afficher les résultats : nombre de vues, de "*like*",

[15] Elsevier est aujourd'hui le premier contributeur à RePEc.

[16] Voir par exemple : ideas.repec.org/top/ (consultée le 03.10.2012).

de téléchargements, de "*followers*", d'amis académiques… Ces possibilités inédites ont donné lieu à ce que certains considèrent comme une véritable explosion des mesures de différentes productions scientifiques (Van Noorden 2010). Elles ont ainsi ouvert la voie à une conception évaluative qui s'oppose à la sélection de quelques outils supposément pertinents et robustes, au profit de l'agrégation d'une pluralité toujours plus importante d'indices[17]. Avec la montée de ces mesures alternatives, intimement liée au mouvement de l'open access, émerge l'idée de la possible appréhension d'un "*total impact*" d'une large gamme de productions scientifiques[18]. Dans ce cadre, les limites traditionnelles de mesures de l'usage par d'autres productions sont totalement effacées : tout lecteur ou "citeur" via des réseaux sociaux, même non académique, est pris en compte.

La création de *PloS One* en 2006 constitue une manifestation exemplaire de ce mouvement. Publiée par l'organisation à but non lucratif Public Library of Science (*PloS*), qui promeut la diffusion gratuite des productions scientifiques, cette revue à vocation généraliste se présente comme l'expression d'un nouveau genre. Tout d'abord, les articles sont non seulement publiés au fil de l'eau, mais ils sont sous licence libre, les auteurs conservant en contrepartie la propriété intellectuelle de leur texte et ayant à leur charge les frais de publication. Ensuite, l'évaluation par les pairs est quelque peu aménagée : tous les articles considérés comme "techniquement justes" sont publiés, et les auteurs des manuscrits rejetés ont un recours par une procédure d'appel formelle. Enfin, depuis mars 2009, chaque article est accompagné sur sa page web d'une palette d'indices, intitulée "*Article Level Metrics*" (ALM), mise à jour en temps réel. Parmi les ALM figurent le nombre de vues de la version HTML, le nombre de téléchargements du fichier PDF, le nombre de citations dans Scopus, WoS, PubMed Central, et le nombre d'occurrences sur CiteUlike, Mendeley, Twitter et FaceBook. Les lecteurs peuvent également noter de une à cinq étoiles la pertinence d'un article, sa fiabilité et son style. L'ensemble de ces mesures est à la fois public et disponible sous la forme d'un fichier XML dont l'usage est à la discrétion de ceux qui le téléchargent.

> "*Article Level Metrics* inscrit des informations transparentes et exhaustives sur l'usage et la portée des articles publiés sur les articles eux-mêmes, afin que l'ensemble de la communauté académique puisse juger de leur valeur"[19].

Pour *PloS One*, la "communauté académique" est donc à la fois composée de producteurs/citeurs classiques pris en compte dans de multiples jeux de données (ISI, Scopus, Google Scholar…), de bloggeurs ou journalistes qui se font relayeurs d'informations (via Twitter, Facebook…), de simples lecteurs d'articles sans être nécessairement des professionnels du monde scientifique, et de téléchargeurs ou stockeurs de données (Mendeley). Mais, comme l'indique le site web dédié aux ALM[20], ces mesures s'adressent également aux institutions de recherche et aux financeurs. Mettant en équivalence l'impact d'une recherche et sa dissémination, le groupe *PloS* soutient que cette dernière ne peut être mesurée qu'ex post, au niveau de l'article.

Pourtant, la revue *PloS One* a été dans le même temps incluse par Thomson dans le WoS en 2010, en reconnaissance de son statut de "haute qualité scientifique", le nombre total de citations que *PloS One* avait reçues de la part des revues indexées dans le SCI étant très élevé[21]. En faisant ainsi partie de celles qui disposent d'un JIF, elle se rapproche des revues plus classiques, qu'elles soient publiées sous forme papier ou seulement en ligne. Bien que

---

[17] Une partie de ces acteurs se sont retrouvés dans le manifeste Altmetrics écrit en 2011, et les conférences qui s'en ont suivi. http://altmetrics.org/manifesto/ (consultée le 30.06.2012).

[18] http://total-impact.org/ (consultée le 30.06.2012).

[19] http://blogs.plos.org/mfenner/2012/04/23/plos-article-level-metrics-interview-with-martin-fenner/ (consultée le 30.06.2012).

[20] http://article-level-metrics.plos.org/ (consultée le 03.10.2012).

[21] http://wokinfo.com/news/plosone/ (consultée le 30.06.2012).

l'article demeure la pierre angulaire sur laquelle sont bâtis les algorithmes et les jeux de données, l'ambivalence est donc flagrante (Davis 2011). D'un côté, la revue est intégrée au monde standard de la bibliométrie, où l'inclusion dans le JCR est à la fois une marque de consécration et un signal de qualité incontournable. De l'autre, elle porte aux nues la multiplication des indices sur chaque article (ALM), au nom de la diffusion qui est une qualité centrale du mouvement de l'open access[22]. La valeur de chaque article est ainsi placée dans les mains des usagers plutôt que dans un processus de peer-review très sélectif en amont de la publication. Et, rappelons-le, ces usagers sont potentiellement très différents de ceux pris en compte dans la fabrication du JCR.

Cette troisième configuration émergente est donc caractérisée par une multitude de jeux de données ouverts qui renvoient à des conceptions d'usagers fortement contrastées, non hiérarchisées entre elles. Du point de vue des algorithmes, on peut dégager une double tendance : d'une part, le recyclage et le raffinement des outils déjà présents dans les deux premières configurations et, d'autre part le comptage brut qui est incorporé dans de nouveaux outils bibliométriques (e.g. nombre de tweets). Parmi ces derniers, l'outil le plus partagé est sans doute le "*top download*" que l'on retrouve chez les acteurs centraux de cette troisième configuration, mais également chez des acteurs plus traditionnels à l'image des "*top 25 hottest articles*" présentés par Elsevier pour chaque champ disciplinaire. De plus, le "*top download*" possède la particularité de se décliner pour toute entité (auteurs, revues, institutions…) et sur tout jeu de données électroniques incluant des documents.

Conclusion

L'attention aux technologies de calcul des productions scientifiques a permis d'identifier trois configurations qui apparaissent au fil du temps. La première se caractérise par la difficile construction d'une position monopolistique par un unique producteur de données, l'ISI, qui a développé plusieurs algorithmes et produit un outil bibliométrique, le JCR, portant sur les revues d'une multitude de domaines scientifiques. La deuxième configuration est marquée par l'apparition de producteurs de données concurrents, jouant sur une plus grande ouverture des données et une plus large couverture des productions scientifiques. C'est ouverture a joué un rôle crucial dans la diffusion d'une myriade d'algorithmes, certains d'entre eux étant stabilisés sous la forme d'outils bibliométriques accessibles à un grand nombre d'acteurs du monde scientifique. À l'image du h-index, ces algorithmes peuvent désormais porter sur les auteurs, les articles, les revues ou les institutions, et instaurent ainsi une véritable banalisation des métriques de la recherche. Enfin la troisième configuration, approfondissant le découplage entre algorithmes, jeux de données et outils bibliométriques, installe un pluralisme des usages et usagers des productions scientifiques. Au cycle de crédibilité scientifique typique de l'intercitation intra-académique (Latour et Woolgar 1979), succèderaient des formes de valorisation plus hétérogènes, construites par des chercheurs simplement lecteurs, ou même par des publics non professionnels qui peuvent télécharger ou signaler un article pour des raisons extra-académiques.

Même si nous avons montré que certains acteurs centraux des deuxième et troisième configurations annonçaient de véritables révolutions bibliométriques, il serait erroné de considérer que l'apparition successive de chaque de ces trois configurations se traduirait par la disparition de la précédente. Au contraire, des producteurs de données ou d'outils déjà présents s'approprient des algorithmes inédits (à l'instar de l'*influence weight*, du h-index et de "*top download*"), et les nouveaux producteurs intègrent, au moins pour partie, les algorithmes (e.g. JIF et *influence weight* dans RePEc) et les jeux de données existants (e.g. Scopus, Google Scholar, PubMed Central pour ALM). De la sorte, les trois configurations

[22] http://blogs.plos.org/plos/2010/01/plos-one-indexed-by-web-of-science/ (consultée le 03.10.2012).

coexistent : non seulement les outils sont simultanément disponibles, mais les mêmes entités sont aussi prises en compte par différentes métriques, à l'image de *PloS One* qui défend des outils alternatifs tout en étant indexée dans le WoS.

Face à la diversification et la diffusion des technologies de calcul dédiées à l'évaluation quantitative des productions scientifiques, est-il encore pertinent de les rassembler sous un vocable unique ? Est-on face à "la" bibliométrie ou faut-il, au contraire, distinguer "des" bibliométries ? Lorsque l'attention des concepteurs ou des utilisateurs d'outils se focalise sur un seul des trois composants que nous avons déplié au fil de cet article, ils privilégient généralement la vision d'un tout unifié et cohérent dépendant uniquement de ce composant précis. Ainsi, quand J. Hirsch veut subsumer la carrière d'un chercheur sous un seul nombre (le h-index), seul le bon algorithme compte, et non le jeu de données pertinents ou l'outil qui permettrait son opérationnalisation. Symétriquement, lorsqu'Elsevier présente son jeu de données Scopus comme étant le plus pertinent et le plus étendu pour évaluer l'impact des recherches, la question des algorithmes ou des outils est secondaire. Du côté des usagers, certains domaines, comme les sciences biomédicales, demeurent fortement attachés à l'outil phare de la première configuration, à l'instar d'un EPST français où l'on considère qu'un fait scientifique marquant consiste en une publication dans une revue dont le JIF est supérieur ou égal à 20. Ailleurs, des chercheurs ou des institutions utilisent le h-index affiché par PoP comme base de négociation pour la rémunération. Dans de telles perspectives, la réponse à la question est singulière : on est clairement dans "la" bibliométrie.

Même si les producteurs d'algorithmes, de jeux de données ou d'outils vantent leur bibliométrie comme étant, si ce n'est la seule possible, au moins la plus fiable et la plus adéquate à ce que l'on souhaite mesurer, il existe en revanche aujourd'hui, du point de vue de l'offre marchande et non marchande, une multitude de bibliométries (Van Noorden 2010). Cela engendre des situations où les conventions permettant de s'accorder sur ce qui fait la qualité des entités évaluées (articles, revues, institutions, chercheurs) sont loin d'être partagées. La manufacture bibliométrique est le lieu privilégié de discussions sur le périmètre des jeux de données, de débats sur la justesse des algorithmes, voire de disputes sanglantes sur la portée et les limites des différents outils bibliométriques disponibles[23]. Comme toute production statistique, les résultats engendrés par ces différentes technologies de calcul ne sont pas toujours considérés comme réalistes par leurs concepteurs (Desrosières 1993). Lorsque des scientomètres testent comparativement les jeux de données ISI, Scopus et Scholar, ils mettent à l'épreuve leur robustesse, leur extension, ou leur pertinence selon les types de questions posées. Quand travaillant autour du h-index, on cherche à tenir compte de la contribution des coauteurs (Schreiber 2009) ou à déterminer sa robustesse selon les disciplines, on teste les limites d'un algorithme ou on en produit plusieurs variantes spécialisées. Comme on l'a vu avec le JCR et les ALM, la plupart des outils bibliométriques stabilisés contiennent eux-mêmes une variété d'algorithmes ou de jeux de données, les utilisateurs faisant face à une pluralité d'options, de variantes et de réglages. Dans chaque cas, la dimension relationnelle entre les trois composants des infrastructures informationnelles élaborées pour l'évaluation quantitative de la recherche est centrale, construisant "des" bibliométries. La troisième configuration en émergence encourage cette vision relationnelle, en rendant simultanément disponible une multitude d'outils non hiérarchisés entre eux.

---

[23] Voir le numéro spécial de *Scientometrics* (2012, vol. 92, n°2), consacré aux débats sur le JIF, où l'un des résumés affirme qu'un des contributeurs à ce numéro montre "tellement de démagogie, d'ignorance et d'arrogance, tant de préjugés et de partialité, tant d'erreurs profondes dans l'utilisation des bases de données, des algorithmes, et dans l'interprétation des résultats qu'il est très peu probable que [ses] articles soient conçus comme de réelles contributions scientifiques".

Remerciements

Cet article n'aurait pas vu le jour sans les encouragements de Jérôme Denis (Telecom ParisTech) qui nous ont incités à l'écrire, l'enthousiasme que nous a communiqué Dominique Cardon (Orange Labs) tout au long de son écriture, et la relecture minutieuse et rigoureuse d'Anne Bertrand (EHESS) à laquelle doit beaucoup la version finale de cet article. Nous remercions également les évaluateurs de la revue *Réseaux* pour leurs suggestions de réaménagement d'une précédente version.

Références

AIUTI F., BARONI C., CAO A. et FANTONI A., 1991, « Academic promotion in Italy », *The Lancet*, vol. 338, n° 8778, p. 1337.

ALMIND T., 1997, « Informetric analyses on the World Wide Web: methodological approaches to 'Webometrics' », *Journal of Documentation*, vol. 53, n° 4, p. 404-426.

Annual Review of Phytopathology, 1975, « Preface », *Annual Review of Phytopathology*, vol. 13, n° 1.

ARCHAMBAULT É. et LARIVIÈRE V., 2009, « History of the journal impact factor: Contingencies and consequences », *Scientometrics*, vol. 79, n° 3, p. 635-649.

AURANEN O. et NIEMINEN M., 2010, « University research funding and publication performance – An international comparison », *Research Policy*, vol. 39, p. 822-834.

BAR-ILAN J., 2007, « Which h-index? — A comparison of WoS, Scopus and Google Scholar », *Scientometrics*, vol. 74, n° 2, p. 257-271.

BAR-ILAN J., 2010, « Citations to the "Introduction to informetrics" indexed by WOS, Scopus and Google Scholar », *Scientometrics*, vol. 82, n° 3, p. 495-506.

BAZERMAN C., 1988, *Shaping written knowledge: the genre and activity of the experimental article in science*, Madison, Wisconsin, The University of Wisconsin Press.

BENSMAN S.J., 2007, « Garfield and the Impact Factor: the creation, utilization and validation of a citation measure, Part 2. », p. http://garfield.library.upenn.edu/bensman/bensmane.

BERGSTROM C., 2007, « Eigenfactor. Measuring the value and prestige of scholarly journals », *College & Research Libraries News*, n° 68, p. 314-316.

BOLLEN J., HERBERT, HAGBERG A. et CHUTE R., 2009, « A principal component analysis of 39 scientific impact measures », *PloS One*, vol. 4, n° 6, p. e6022-e6022.

BOLLEN J., RODRIQUEZ M.A. et VAN DE SOMPEL H., 2006, « Journal status », *Scientometrics*, vol. 69, n° 3, p. 669-687.

BOWKER G.C. et STAR S.L., 1999, *Sorting things out. Classification and its consequences*, Cambridge, MA, MIT Press.

BRADFORD, 1934, « Sources of Information on Specific Subjects », *Engineering: An Illustrated Weekly Journal*, vol. 137, n° 3550, p. 85-86.

BRAUN T., GLÄNZEL W. et SCHUBERT A., 2006, « A Hirsh-type index for journals », *Scientometrics*, vol. 69, n° 1, p. 169-173.

BROOKS T.A., 1986, « Evidence of complex citer motivations », *Journal of the American Society for Information Science*, vol. 37, n° 1, p. 34-36.

BURNHAM J.F., 2006, « Scopus database: a review », *Biomedical Digital Libraries*, vol. 3, n° 1, p. 1-8.

CARDON D., 2013, « Dans l'esprit du PageRank. Une enquête sur l'algorithme de Google », *Réseaux*, n°177, p. 63-95.

COLE J.R. et COLE S., 1971, « Measuring the quality of sociological research: problems in the use of the Science Citation Index », *American Sociologist*, vol. 7, n° 1, p. 23-29.

DAVIS P.M., 2011, « PLoS One's 2010 Impact Factor », *The Scholarly Kitchen*, http://scholarlykitchen.sspnet.org/2011/06/28/*plos*.

DESROSIÈRES A., 1993, *La Politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte.

EGGHE L., 2005, « Expansion of the field of informetrics: Origins and consequences », *Information Processing & Management*, vol. 41, n° 6, p. 1311-1316.

EGGHE L., 2006, « Theory and practise of the g-index », *Scientometrics*, vol. 69, n° 1, p. 131-152.

FABBRI L.M., 1987, « Rank injustice and academic promotion », *The Lancet,* vol. 330, n° 8563, p. 860.

FALAGAS M.E., PITSOUNI E.I., MALIETZIS G.A. et PAPPAS G., 2008, « Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses », *FASEB journal*, vol. 22, n ° 2, p. 338-342.

FERRARA E., 2012, « V-index: a novel metric to measure virtuosity of academics », p. Arxiv preprint arXiv:1202.3119.

GARDNER S. et ENG S., 2005, « Gaga over Google? Scholar in the Social Sciences », *Library Hi Tech News*, vol. 22, n° 8, p. 42-45.

GARFIELD E., 1972, « Citation analysis as tool in journal evaluation », *Science*, vol. 178, n° 4060, p. 471-479.

GARFIELD E., 1976, « Significant journals of science », *Nature*, vol. 264, n° 5587, p. 609-615.

GARFIELD E., 1977, « Will ISI's Arts & Humanities Citation Index Revolutionize Scholarship », *Current Contents*, vol. 32, p. 204-208.

GARFIELD E., 1979, « Is citation analysis a legitimate evaluation tool? », *Scientometrics*, vol. 1, n° 4, p. 359-375.

GARFIELD E. et SHER I.H., 1963, « New Factors in the Evaluation of Scientific Literature Through Citation Indexing », *American Documentation*, vol. 14, n° 3, p. 195-201.

GILES J., 2005, « Science in the web age: start your engines », *Nature*, vol. 438, n° 7068, p. 554-555.

GINGRAS Y., 2008, « Du mauvais usage de faux indicateurs », *Revue d'histoire moderne et contemporaine*, vol. 55, n° 4bis, p. 67-79.

GROSS P.L.K. et GROSS E.M., 1927, « College Libraries and Chemical Education », *Science*, vol. 66, n° 1713, p. 385-389.

GUNNARSDÓTTIR K., 2005, « Scientific journal publications: on the role of electronic preprint exchange in the distribution of scientific literature », *Social Studies of Science*, vol. 35, n° 4, p. 549-579.

HIRSCH J.E., 2005, « An index to quantify an individual's scientific research output », *Proceedings of the National Academy of Sciences*, vol. 102, n° 46, p. 16569-16572.

HORNE R., PETRIE K. et WESSELY S., 2009, « H-index pathology », *British Medical Journal*, vol. 339, n° 7735, p. 1447-1448.

IGLESIAS J.E. et PECHARROMÁN C., 2007, « Scaling the h-index for different scientific ISI fields », *Scientometrics*, vol. 73, n° 3, p. 303-320.

JACSO P., 2004, « Scopus », Peter's Digital Reference Shelf, Thompson-Gale web site.

KARLSSON S. et KRICHEL T., 1999, RePEc and S-WoPEc : Internet access to electronic preprints in economics, *Electronic publishing' 99*. Ronneby.

KLEINBERG J.M., 1999, « Authoritative sources in a hyperlinked environment », *Journal of the ACM*, vol. 46, n° 5, p. 604-632.

KOSMULSKI M., 2011, « Successful papers: a new idea in evaluation of scientific output », *Journal of Informetrics*, vol. 5, n° 3, p. 481-485.

LATOUR B., 1989, *La Science en action*, Paris, La Découverte.

LATOUR B. et WOOLGAR S., 1979, *Laboratory life: the social construction of scientific facts*, Beverly Hills, Sage.

LAVIE P., 2009, « The race for the impact factor », *Journal of Sleep Research*, vol. 18, n° 3, p. 283-284.

LIU M., 1993, « Progress in documentation the complexities of citation practice: a review of citation studies », *Journal of Documentation*, vol. 49, n° 4, p. 370-408.

MACDONALD S. et KAM J., 2007, « Aardvark et al.: quality journals and gamesmanship in management studies », *Journal of Information Science*, vol. 33, n° 6, p. 702-717.

MEHO L.I. et YANG K., 2006, « A New Era in Citation and Bibliometric Analyses: Web of Science, Scopus, and Google Scholar », p. 49.

MORAVCSIK M.J., 1985, *Strengthening the coverage of the third world science*, National Science Foundation.

MOYA-ANEGÓN F., CHINCHILLA-RODRÍGUEZ Z., VARGAS-QUESADA B., CORERA-ÁLVAREZ E., MUÑOZ-FERNÁNDEZ F.J., GONZÁLEZ-MOLINA A. et HERRERO-SOLANA V., 2007, « Coverage analysis of Scopus: A journal metric approach », *Scientometrics*, vol. 73, n° 1, p. 53-78.

NARIN F., 1976, *Evaluative bibliometrics: the use of publication and citation analysis in the evaluation of scientific activity*, Cherry Hill, New-Jersey, Computer Horizons.

NARIN F., 1994, « Patent bibliometrics », *Scientometrics*, vol. 30, n° 1, p. 147-155.

NORRIS M. et OPPENHEIM C., 2007, « Comparing alternatives to the Web of Science for coverage of the social sciences literature », *Journal of Informetrics*, vol. 1, n° 2, p. 161-169.

OPTHOF T., 1997, « Sense and nonsense about the impact factor », *Cardiovascular research*, vol. 33, n° 1, p. 1-7.

PAGE L., BRIN S., MOTWANI R. et WINOGRAD T., 1999, « The Pagerank citation ranking: bringing order to the web », Technical Report, Stanford InfoLab, p. 1-17.

PINSKI G. et NARIN F., 1976, « Citation Influence for Journal Aggregates of Scientific Publications: Theory, with Application to the Literature of Physics », *Information Processing and Management*, vol. 12, n° 5, p. 297-312.

PONS-NOVELL J. et TIRADO-FABREGAT D., 2010, « Is there life beyond the ISI Journal lists? The international impact of Spanish, Italian, French and German economics journals », *Applied Economics*, vol. 42, n° 6, p. 689-699.

PONTILLE D., 2004, *La signature scientifique. Une sociologie pragmatique de l'attribution*, Paris, CNRS Editions.

PONTILLE D. et TORNY D., 2010, « Revues qui comptent, revues qu'on compte : produire des classements en économie et gestion », *Revue de la Régulation*, n° 8, p. [en ligne].

PRIME C., BASSECOULARD E. et Zitt M., 2002, « Co-citations and co-sitations : A cautionary view on an analogy », *Scientometrics*, vol. 54, n° 2, p. 291-308.

RAISIG L.M., 1960, « Mathematical Evaluation of the Scientific Serial », *Science*, vol. 131, n° 3411, p. 1417-1419.

SCHREIBER M., 2009, « A case study of the modified Hirsch index hm accounting for multiple coauthors », *Journal of the American Society for Information Science and Technology*, vol. 60, n° 6, p. 1274-1282.

SEGLEN P.O., 1997, « Why the impact factor of journals should not be used for evaluating research », *British Medical Journal*, vol. 314, n° 7079, p. 498-502.

SOLLA PRICE D.J., 1963, *Little Science, Big Science*, New York, Columbia University Press.

STAR S.L., 1995, The politics of formal representations: wizards, gurus, and organizational complexity, in S.L. Star (ed.), *Ecologies of Knowledge. Work and Politics in Science and Technology*, Albany, NY: SUNY Press.

TAUBES G., 1993, « Measure for measure in science », *Science*, vol. 260, n° 5110, p. 884-886.

VAN NOORDEN R., 2010, « Metrics: a profusion of measures », *Nature*, vol. 465, n° 7300, p. 864-866.

VAN RAAN A.F.J., 2005, « Fatal attraction : conceptual and methodological problems in the ranking of universities by bibliometric methods », *Scientometrics*, vol. 62, n° 1, p. 133-143.

VAN RAAN A.F.J., 2006, « Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups », *Scientometrics*, vol. 67, n° 3, p. 491-502.

WOUTERS P., 1999, *The Citation Culture*, Ph.D. Thesis, Amsterdam, University of Amsterdam.

WOUTERS P., 2006, « Aux origines de la scientométrie. La naissance du Science Citation Index », *Actes de la Recherche en Sciences Sociales*, n° 164, p. 10-22.