

Think small! Science disruptive et petites équipes

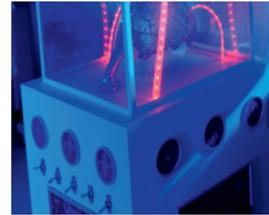
Jacques Haiech¹, Thierry Jouault^{2,3}

La bibliométrie¹ et la scientométrie² sont des aides à la décision pour les politiques de recherche et d'enseignement supérieur. Le Graal des scientométristes est ainsi d'être capables de distinguer entre une recherche disruptive, qui ouvre de nouveaux territoires de la connaissance, et une recherche confirmatoire, qui laboure ces nouveaux lieux avec force, en leur apportant de possibles développements scientifiques et techniques.

Pour l'atteindre, les scientométristes doivent donc être capables de distinguer dans un corpus d'articles ou de brevets³, ceux qui relèvent d'une recherche disruptive de ceux relatant les résultats d'une recherche confirmatoire.

Un nouvel indicateur bibliométrique

De nombreux indicateurs existent. Ils permettent une évaluation du niveau des publications réalisées par un chercheur ou une équipe en fonction du facteur d'impact de la revue, de la place du chercheur dans la liste des auteurs, ce que réalise, par exemple, le score SIGAPS/SAMPRA⁴. Ils permettent aussi d'évaluer la notoriété du chercheur au travers de la mesure des citations de ses publications par les laboratoires qui y font référence (avec *Publons*⁵ ou le *Web of Science*⁶). C'est le rôle du facteur *h*⁷ et de ses variantes (facteur *g* d'Egghe). Mais aucun ne permet de distinguer les deux types de recherche, disruptive et confirmatoire, et



¹CNRS UMR7242 BSC, ESBS, 300 Bd Sébastien Brant, CS 10413, 67412 Illkirch cedex, France.

²Univ. Lille, Inserm, CHU Lille, U1286 - Infinite - Institut de recherche translationnelle en Inflammation, F-59000 Lille, France.

³Rédacteur en chef adjoint de *médecine/sciences*.

haiech@hotmail.fr

si l'article cité (et donc son auteur/ses auteurs) relate une recherche originale

influençant un paradigme, ou s'il s'agit d'une revue de la littérature que toute la communauté citera. Dans les deux cas, le taux de citations sera important, mais les conséquences sur les recherches ultérieures seront fondamentalement différentes.

Pour construire un indicateur capable de réaliser cette distinction, deux stratégies sont possibles. La première est dictée par les données accessibles. Il est possible de construire un corpus d'articles contenant des articles relatant une recherche disruptive (par exemple les articles qui ont justifié qu'un auteur/une auteure obtienne un prix Nobel *versus* des articles très bien cités mais qui en fait ne participent pas à la création d'un nouveau paradigme). Des techniques d'apprentissage des règles permettant de distinguer entre ces deux sous-ensembles d'articles sont alors mises en place et on utilise ensuite ces règles pour classer l'ensemble des articles publiés. Une autre possibilité consiste à prendre en compte le temps dans l'analyse. Plus le temps passe, moins l'article sera cité, sauf s'il a contribué réellement à l'avancée de la connaissance et est toujours d'actualité, le facteur d'impact de la revue dans laquelle il a été publié ne sem-

(→) Voir le Dossier technique de L. Baudoin et al., *m/s* octobre 2004, page 909

blant pas un critère de longévité [1] (→) L'autre stratégie est fondée sur l'établissement d'un modèle. C'est celle qu'ont présentée Wu *et al.* dans un article publié dans la revue *Nature* [2], commenté par P Azoulay [3]. Ils décrivent et valident un indice reposant sur les citations, qui avait été proposé précédemment pour les brevets [4].

L'hypothèse des auteurs est de définir la nature de l'article par les citations qu'il engendre : citation de l'article lui-même ou citation des références de cet article. Considérons un article X publié par *John Doe* et l'ensemble Y des articles qu'il cite en référence. Si cet article est cité par des publications Z qui ne font référence à aucun ou peu d'articles de l'ensemble Y, alors il sera considéré comme un article relatant une recherche disruptive, c'est-à-dire un article original et perturbateur. Inversement, si l'article Z, qui cite

Vignette (Photo © Inserm - François Guénet).

¹ L'Inserm, par exemple, possède une cellule « mesures, indicateurs et bibliométrie » qui suit les activités de l'organisme. Cette cellule réalise les indicateurs de production académique pour l'évaluation et le pilotage. Elle fait partie du département « Évaluation et suivi des programmes » de cet organisme. <https://www.inserm.fr/professionnels-recherche/bibliometrie>

² La scientométrie est un domaine d'étude lié à la bibliométrie et à la cybermétrie qui vise à analyser les publications scientifiques.

³ Un brevet protège une invention qui doit être nouvelle, marquée par une activité inventive, et susceptible d'applications industrielles. Il peut être issu d'une recherche disruptive ou d'une recherche confirmatoire. Dans ce dernier cas, il peut plus facilement faire l'objet de contestations et de batailles juridiques.

⁴ *software for analysis and management of publications and research assessment*.

⁵ Publons utilise les publications directement importées du Web of Sciences et d'ORCID, ainsi que les données de citations du « Web of Science Core Collection ».

⁶ Le Web of Science est géré actuellement par la firme privée américaine (dont le siège est à Philadelphie) Clarivate Analytics qui détient également Publons et le Journal of Citation Reports qui publie le facteur d'impact de très nombreuses revues.

⁷ Le facteur *h* (ou *h-index*), créé par le physicien Jorge Hirsch en 2005, est un indicateur d'impact des publications d'un chercheur ou d'une équipe. Il prend en compte le nombre de publications et le nombre de leurs citations.



Doe *et al.*, cite un grand nombre de références de l'ensemble Y, l'article princeps de Doe *et al.* sera alors considéré comme relatant une recherche confirmatoire, reprenant des données originales (fournies par l'ensemble Y) et les confirmant ou les synthétisant. C'est sur cette idée que Wu *et al.* ont construit un indicateur prenant une valeur qui sera comprise entre -1 (correspondant à une recherche confirmatoire) et +1 (recherche disruptive) (voir *Encadré* ci-dessous).

Construction de l'indicateur D pour l'article X

Considérons un article X. Celui-ci peut être cité par un article Y et celui-ci ne citera pas les références citées dans l'article X. L'article Z sera dans l'ensemble i.

L'article Y cite l'article X et cite aussi les références de l'article X. Il sera dans l'ensemble j.

Enfin, l'article Y ne cite pas l'article X mais cite les références de l'article X. Il sera dans l'ensemble k.

Soit n_i , n_j et n_k le nombre d'articles de chaque ensemble i, j, k, l'indicateur D pour l'article X est défini par :

$$D = \frac{n_i - n_j}{n_i + n_j + n_k}$$

Wu *et al.* ont validé cet indicateur sur un corpus d'articles, comme défini dans la première stratégie. Ils l'ont utilisé pour associer aux articles référencés dans le *Web of Science* entre 1900 et 2014 (soit près de 26 millions d'articles), le nombre D (compris entre -1 et 1) illustrant le niveau de recherche disruptive.

Intuitivement, on considère qu'un article relatant une recherche disruptive ($D = 1$) se détachera des articles qu'il cite en référence et sera donc cité pour lui-même car considéré par le lecteur comme suffisamment perturbateur du paradigme : lorsqu'il sera cité, il ne sera pas associé aux articles qu'il a cités puisque le travail relaté ne sera pas une confirmation d'un travail précédent qui serait dans les références. À titre d'exemple, le score de disruption obtenu par l'article de J.D. Watson et F.H. Crick, paru dans *Nature* en 1953 [5], est de 0,88, proche de 1. Il s'agit donc d'un article hautement cité qui entre dans la catégorie des articles perturbateurs. Le nombre de citations de cet article⁸ était, en 2019, de 6 380, d'après *Nature*⁹ et de 12 809 d'après *Google Scholar*¹⁰, ce dernier étant plus enclin à compter les citations par tout type d'article.

Un lien entre taille des équipes et effet perturbateur

Armés de cette nouvelle mesure, Wu *et al.* ont analysé l'influence de la taille de l'équipe publiante sur l'aspect novateur de ses publications. Ils montrent que le type de travail effectué par les grandes équipes et les petites équipes diffère de manière marquée : les petites

équipes sont beaucoup plus susceptibles de publier des articles perturbateurs que les grandes équipes. Cette constatation est valable pour les articles, comme pour les brevets. Cela est également vérifié pour toutes les disciplines scientifiques, de la biologie aux sciences physiques, en passant par les sciences humaines et sociales.

Ainsi, les petites équipes feraient de la recherche disruptive et les gros consortiums de la recherche confirmatoire, en confirmant et incrémentant, parfois tardivement, les résultats précédemment obtenus par des petites équipes. En utilisant les données associées aux articles permettant de connaître le mode de financement de la recherche réalisée dans l'article, les auteurs montrent que la recherche financée sur projet serait plutôt confirmatoire. Ce type de financement a eu tendance à s'accroître dans la dernière décennie, entraînant même les petites équipes à se tourner vers la recherche confirmatoire afin d'obtenir ces financements thématiques.

Leur analyse d'un échantillon important d'environ 38 millions de chercheurs et de leurs travaux publiés montre que les scientifiques participent à un plus grand nombre de projets de consolidation quand ils font partie d'une grande équipe que lorsqu'ils travaillent en petite équipe.

L'article de Wu *et al.* laisse ouverte la question des mécanismes : pourquoi les petites équipes seraient-elles plus susceptibles d'effectuer des travaux perturbateurs ? Quel est le chevauchement des compétences, des antécédents et de l'expérience des membres des petites équipes et de ceux des grandes équipes ? Les différences de talent entre ces chercheurs sont-elles plus ou moins prononcées dans les petites équipes scientifiques que dans les grandes collaborations ? Ces questions nécessitent un examen plus approfondi.

Conclusion

Les conclusions des auteurs sont fondées sur des corrélations et ne sont pas nécessairement des relations causes-effets. Des variables cachées pourraient en effet exister ou des modèles plus complexes que ces simples modèles linéaires. Si les conclusions se confirment néanmoins, certaines recommandations pour nos politiques de recherche pourraient être faites, en particulier : éviter de ne financer la recherche que sur des appels d'offres afin de maintenir un vivier de petites équipes, aptes à réaliser les projets remettant en question les paradigmes, mais pour lesquelles les résultats de ces appels sont souvent un échec ; favoriser l'interaction ou le maillage entre ces petites équipes et les

⁸ Ces analyses de citations ne tiennent pas compte des compétitions géopolitiques entre le monde anglo-saxon, qui contrôle les revues scientifiques, et le reste du monde, avec, parfois, la rétention d'articles soumis jusqu'à ce qu'une équipe anglo-saxonne soumette un travail similaire.

⁹ Source : *Nature.com*.

¹⁰ <https://scholar.google.co.uk/citations?user=fZQhpvkAAAAJ&hl=en>

grands consortiums, afin de permettre à ces derniers d'étayer leurs recherches sur celles, plus disruptives, des petites équipes. Il pourrait être également intéressant de définir un mécanisme récompensant les petites équipes (et/ou leurs universités) lorsque leurs idées sont utilisées par des grands consortiums qui, eux, réussiront à lever les fonds des appels à projets. Cela permettrait de mettre en place un cercle vertueux entre petites équipes et grands consortiums. ◊

Think small! Disruptive science comes from small research teams

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Baudoin L, Haeffner-Cavaillon N, Pinhas N. Indicateurs bibliométriques : réalités, mythes et prospective. *Med Sci (Paris)* 2004 ; 20 : 909-15.
2. Wu L, Wang D, Evans JA. Large teams develop science and technology; small teams disrupt it. *Nature* 2019 ; 566 : 378-82.
3. Azoulay P. Small research teams disrupt science more radically than large ones. *Nature* 2019 ; 566 : 330-2.
4. Funk RJ, Owen-Smith J. A dynamic network measure of technological change. *Management Science* 2017 ; 63 : 791-817.
5. Watson JD, Crick FH. Molecular structure of nucleic acids. *Nature* 1953 ; 171 : 737-8.

TIRÉS À PART

J. Haiech

m/s
médecine/sciences

/ Lieu privilégié d'écriture pour des chercheurs et cliniciens-chercheurs reconnus au plan international..., mais aussi pour les jeunes scientifiques et médecins en formation.

/ Éditoriaux, Synthèses, Nouvelles et Dossiers techniques : une information scientifique et médicale intelligible, précise et mise en perspective.

/ Des analyses et débats sur l'Histoire de la médecine et de la biologie, un regard critique sur les retombées diagnostiques et thérapeutiques des dernières avancées et leurs dimensions éthiques dans les Forums, Chroniques et Perspective/Horizons.

/ Un panorama clair et concis de l'actualité scientifique : des schémas et tableaux synthétiques et lisibles, des données chiffrées, une bibliographie utile, afin qu'aucun fait biologique et médical ne vous échappe.

médecine/sciences est indexé dans PubMed/Medline, Current Contents, série Life Sciences, EMBASE/Excerpta Medica PASCAL CABS BIOSIS

Avec *m/s*, vivez en direct les progrès et débats de la biologie et de la médecine

CHAQUE MOIS / AVEC LES ARTICLES DE RÉFÉRENCE DE *m/s*
CHAQUE JOUR / SUR WWW.MEDECIENSCIENCES.ORG

Pour recevoir un numéro de *m/s* gratuitement inscrivez-vous sur www.medecinesciences.org

