

Les pièges de l'évaluation quantitative de la recherche

Jean-Paul Delahaye

Les systèmes numériques d'évaluation de la recherche sont utiles mais fragiles. Ils sont engagés dans une course au perfectionnement incessante avec les méthodes de fraude qui se développent pour les berner.

La façon dont les scientifiques communiquent entre eux et dont on les évalue change rapidement. La pression exercée sur les chercheurs pour qu'ils publient incite aussi à la multiplication du nombre de publications (*voir la courbe page 55*). Des bases de données d'articles permettent de compter les publications d'un auteur et les références qui y sont faites, ce qui a conduit à imaginer des formules de calcul évaluant automatiquement la qualité des chercheurs, formules dont il faut connaître les subtilités, les perfectionnements récents et les limites. Si certains outils (avec ou sans système d'IA générative) permettent de créer de faux articles, d'autres heureusement servent à détecter les publications frauduleuses.

Un jeu du chat et de la souris s'est mis en place entre notateurs et notés, car les évaluations numériques jouent un rôle important dans les laboratoires

et les universités pour le recrutement et l'évaluation des scientifiques. L'université de l'État de l'Ohio, parmi bien d'autres, expliquait encore il y a quelques mois l'importance de ces indicateurs et évoquait l'indice h , qui vingt-cinq ans après sa création tient toujours une place centrale: «Dans le monde académique, les indicateurs jouent un rôle crucial dans l'évaluation de l'impact de la recherche et de la productivité des auteurs. L'un des indicateurs les plus reconnus est l'indice h .»

Cet indice (*h-index* en anglais) et ses variantes attribuées aux chercheurs pour mesurer leur valeur, ainsi que les facteurs d'impact et classements des revues scientifiques prétendant déterminer le sérieux de celles-ci, sont des outils délicats qu'un éclairage mathématique aide à mieux saisir, afin d'en cerner les limites et d'identifier les fraudes auxquelles ils exposent les évaluateurs.

L'indice h d'un chercheur ou d'une chercheuse, ou même d'un laboratoire ou d'un groupe étendu de scientifiques est un nombre entier h qui dépend de ses publications et du nombre de fois qu'elles ont été mentionnées – nombre de citations – dans d'autres publications. Cette notation a été conçue en 2005 par Jorge Hirsch, un physicien américain qui voulait disposer d'un moyen rapide d'évaluation du travail de ses collègues. Cet indice est fragile et plus difficile qu'il n'y paraît à calculer précisément et à

L'essentiel

L'évaluation des scientifiques repose de plus en plus sur des indicateurs bibliométriques comme l'indice h , qui mesure à la fois la productivité et l'impact via les citations.

Ces outils, dépendants notamment des bases de données utilisées, présentent des limites méthodologiques et de fiabilité.

Des fraudes complexes biaisent ces indicateurs, allant de l'autopromotion à la manipulation des métadonnées, comme l'a par exemple révélé la méthode des «références furtives».

Des indicateurs complémentaires

Disposant d'une base d'articles, chacun indiquant ses auteurs et les articles qu'il cite, il est possible de définir plusieurs indicateurs numériques mesurant le travail et la valeur de ce dernier pour une entité E qui publie (chercheur, chercheuse, laboratoire, etc.).

- * N: nombre d'articles dont E est signataire.
- * N': nombre d'articles dont E est signataire, en ne comptant que 1/k pour un article ayant k signataires.
- * C: nombre de citations faites aux articles dont E est signataire.
- * C': nombre de citations faites aux articles publiés par E, en divisant par k le compte d'un article s'il a k signataires.
- * L'indice h (proposé par Jorge Hirsh en 2005): plus grand entier h tel que E est signataire de h articles, chacun cité au moins h fois. Cet indice h corrige le défaut de N ou N' qui accordent trop de valeur au travail d'une entité qui publie beaucoup mais n'est que rarement cité. L'indice h corrige aussi le défaut de C et C' qui accordent trop de valeur à une entité qui a publié très peu d'articles (voire un seul) ayant réussi à se faire beaucoup citer.
- * L'indice h' est défini comme h mais en prenant en compte, comme pour C', le nombre d'auteurs.
- * L'indice g (proposé par Leo Egghe en 2006): plus grand entier g tel que le total des citations aux g articles les plus cités de E est au moins égal à g². Il possède les mêmes qualités que l'indice h, qu'il dépasse toujours, mais prend mieux en compte les articles beaucoup cités.

* L'indice g' est défini comme g, mais en prenant en compte le nombre d'auteurs.

* L'indice e (proposé par Şirag Erkol *et al.* en 2022) est défini par la formule ci-dessous, où on note c₁, c₂, ..., c_n le nombre de citations aux n articles dont E est signataire et où on note C le total des c_i.

$$e = \frac{1}{n} \sum_{i=1}^n c_i \log \frac{c_i}{C}$$

On vérifie que e est le produit du nombre de citations moyen par article multiplié par l'entropie de Shannon de la distribution des citations. Cet indice valorise les entités E ayant beaucoup de publications bien citées (comme l'indice h et l'indice g), mais prend mieux en compte la régularité de la qualité des publications de E. Cet indice a été conçu en étudiant les publications des Prix Nobel avec pour but de mieux repérer les scientifiques de premier rang.

Le tableau ci-dessous présente cinq exemples d'auteurs imaginaires ayant chacun publié 10 articles. On les a conçus pour aider à comprendre les différences et les complémentarités entre les outils de mesure.

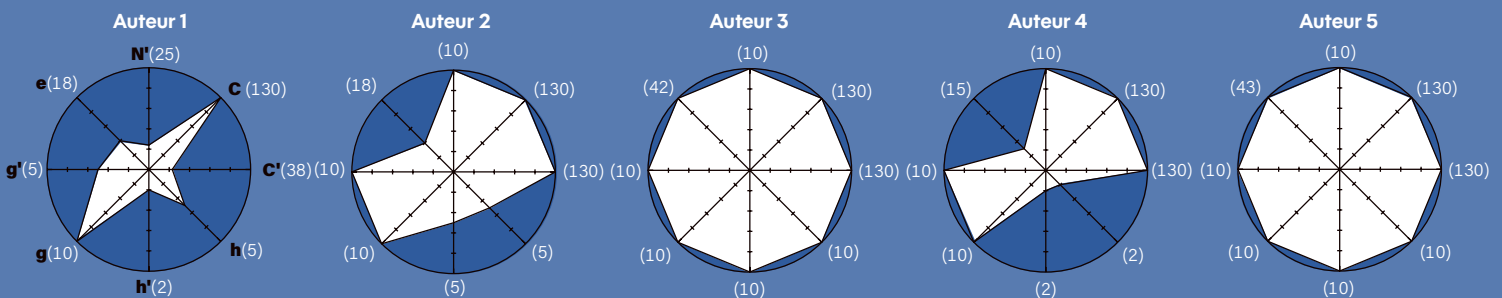
Nos notations signifient que par exemple l'article le plus cité de l'auteur 1 l'a été 100 fois et qu'il avait 5 auteurs, son second article le plus cité l'a été 8 fois et il avait 2 auteurs, etc.

Articles publiés (nombre de citations - nombre d'auteurs)

Auteur 1	100-5	8-2	7-7	5-3	5-2	1-4	1-6	1-8	1-10	1-8
Auteur 2	100-1	8-1	7-1	5-1	5-1	1-1	1-1	1-1	1-1	1-1
Auteur 3	22-1	16-1	15-1	13-1	12-1	11-1	11-1	10-1	10-1	10-1
Auteur 4	100-1	22-1	1-1	1-1	1-1	1-1	1-1	1-1	1-1	1-1
Auteur 5	13-1	13-1	13-1	13-1	13-1	13-1	13-1	13-1	13-1	13-1

Le calcul de N, N', C, C', h, h', g, g' et e donne les profils ci-dessous. Les valeurs obtenues pour chaque indice sont entre parenthèses. Sur chaque axe, elles sont rapportées à la valeur maximum observée dans l'échantillon des cinq chercheurs considérés.

Profils des auteurs



L'auteur 2 a toujours publié seul, a été cité exactement comme l'auteur 1, qui, lui, a toujours publié avec d'autres. Ils ont bien sûr les mêmes coefficients N, C, h, g et e, mais pour N', C', h', g', l'auteur 2 est bien meilleur que l'auteur 1.

L'auteur 2 et l'auteur 3 ont le même nombre de publications et le même nombre total de citations (130). L'auteur 3 a cependant un bien meilleur indice h que l'auteur 2, car il a plus d'articles bien cités. Pour le g, qui ne pénalise pas la concentration des citations, les indices des deux auteurs obtiennent la même note.

Les auteurs 4 et 5 ont publié le même nombre d'articles et obtenu le même nombre total de citations, cependant h est meilleur pour l'auteur 5 que pour l'auteur 4. Là encore c'est le nombre de bons articles qui explique la différence. Pour l'indice e, l'auteur 5 est

sensiblement meilleur que l'auteur 4, mais – et c'est plus remarquable – l'auteur 5 est meilleur aussi que l'auteur 3, alors qu'ils ont les mêmes N, N', C, C', h, h', g, g'. Cela est dû à la répartition parfaite des citations de l'auteur 5 (donc d'entropie log₂(10) = 3,32) alors que celle de l'auteur 3 est plus irrégulière donc d'entropie plus faible.

Tous ces indices et coefficients sont intéressants à la condition d'être calculés soigneusement, ce qui n'est pas facile. Aucun n'est parfait. Le site Google Scholar utilisé à partir de Google Chrome y donne accès, mais les plus grands risques qu'il représente sont l'attribution à un auteur d'articles publiés par un homonyme ou l'oubli d'articles si on n'a pas demandé au site de chercher assez profondément dans les listes dont il dispose.

interpréter. En effet, pour compter les publications et les citations qui y sont faites la base de publications utilisée est déterminante: ce peut être une liste limitée de journaux considérés comme sérieux, ou à l'opposé être une base de données large et tolérante, ce qui augmentera les valeurs calculées. Trois bases aujourd'hui sont principalement utilisées comme ressources pour ces calculs: Web of science, Scopus, Google Scholar. Seule la troisième, la moins stricte, est d'accès gratuit.

Bases tolérantes

Le système de prise en compte des publications et citations étant fixé, l'indice h d'une entité E vaut h , si h est le plus grand entier n tel que E a publié au moins n articles qui ont été cités chacun au moins n fois. Cette double présence de n dans la définition a pour conséquence que si E a publié 50 articles qui n'ont chacun été cités qu'une fois, alors son indice h vaut 1. Extrême inverse: si E a publié un seul article qui a été cité 50 fois, il aura aussi un indice h valant 1. Pour avoir un bon indice h il faut publier beaucoup et être cité beaucoup, par exemple avoir publié 50 articles cités chacun 50 fois ou plus.

L'encadré de la page 53 présente les principaux paramètres et indices alternatifs proposés – depuis longtemps ou tout récemment – en complément à l'indice h . Précisons d'abord que l'interprétation de ces indices varie d'une discipline à l'autre et que de nombreuses sources d'erreurs (homonymie, fautes de frappe, bases incomplètes...) faussent les notes attribuées par ces indices qui ne se fondent pas sur les contenus scientifiques des articles publiés par un auteur, et ne font appel à aucun spécialiste compétent.

Il faut ensuite considérer la possibilité d'une fraude. Un fraudeur (scientifique, équipe, laboratoire...) dispose aujourd'hui de nombreux moyens pour étendre la liste de ses publications et des références faites à son travail: scinder un article long en plusieurs, laisser délibérément des erreurs pour publier ensuite un article de rectification, plagier, etc.

L'existence de revues qualifiées de «prédatrices», qui n'exercent aucun contrôle sur leur contenu en publiant tout ce qu'on leur envoie pourvu que

l'auteur paye les frais exigés pour en donner l'accès libre sur internet a, par ailleurs, aggravé considérablement la difficulté du calcul du nombre de publications véritables d'un auteur E . Des listes de revues prédatrices sont disponibles (voir la bibliographie), et au moindre doute il est conseillé de s'y reporter.

Parmi les plus grosses bêtises qu'on peut observer provenant d'articles publiés dans les revues prédatrices celles qui concernent les mathématiques sont particulièrement déconcertantes et navrantes. Depuis que Ferdinand von Lindemann l'a démontré en 1882, on sait que le nombre π est transcendant, c'est-à-dire qu'il n'est la solution d'aucune équation du type: $a_k X^k + a_{k-1} X^{k-1} + \dots + a_1 X^1 + a_0 = 0$ où les a_i sont des entiers (positifs, nuls ou négatifs). Cela implique en particulier que le problème de la quadrature du cercle «dessiner un carré de la même aire qu'un disque donné, en n'utilisant qu'une règle non graduée et un compas» n'a pas de solution.

Le retour de la quadrature

Cela s'oppose à ce que d'innombrables amateurs et farfelus ont cru en proposant depuis deux mille ans des constructions approximatives ou fausses de la quadrature du cercle. Ce n'était pas très grave, car avant l'ère des revues prédatrices les faux résultats contredisant le théorème de von Lindemann ne passaient pas la barrière des expertises et n'étaient donc pas publiés dans les revues de mathématiques. L'Académie des sciences de Paris avait même adopté une résolution indiquant qu'elle n'examinerait plus les propositions de solutions portant sur ce problème.

Aujourd'hui toute idée, aussi farfelue soit-elle, trouve sa place sur le web et même comme article dans une revue «internationale» prétendument «scientifique». Toute folie est en mesure de se déclarer «validée par la communauté académique», ce qui ne trompe pas si l'on est bien informé, mais peut troubler les amateurs et curieux de science.

J'ai pris conscience de la gravité de cette situation en recevant un livre d'un quarreur de cercle – comme on les nomme parfois – Sarva Jagannadha Reddy, qui se présentait comme «zoologiste, à la retraite» habitant à Tirupati, en Inde. Son livre, écrit en anglais, est intitulé *The Gayatri Pi value* en l'honneur d'une déesse indienne et présente entre autres choses une solution de la quadrature du cercle. Il a ceci de frappant qu'il se compose en fait de 15 articles ayant auparavant été publiés dans des revues scientifiques d'apparence académique, comme l'International Journal of Mathematics and Statistics Invention ou l'International Journal of Latest Trends in Engineering and Technology.

L'AUTEUR



Jean-Paul Delahaye

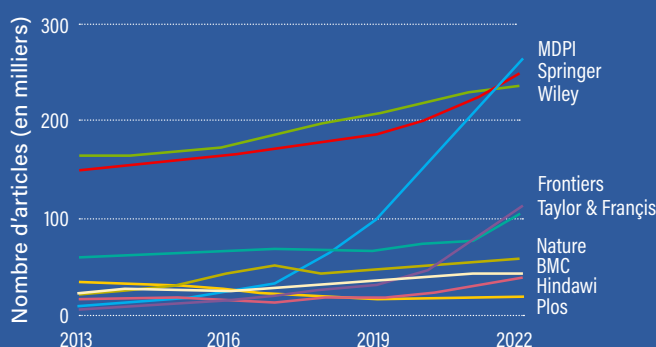
professeur émérite à l'université de Lille et chercheur au laboratoire Cristal (Centre de recherche en informatique, signal et automatique de Lille)

Ce sont bien sûr des revues prédatrices. Qu'une démonstration soit publiée par ignorance ou erreur dans une revue frauduleuse ne signifie pas qu'elle soit fautive; j'ai donc examiné la preuve de Reddy affirmant que $\pi = (14 - \sqrt{2})/4$. Cette valeur est impossible puisque π est transcendant et ne peut donc avoir d'expression de ce type. Le calcul de la valeur proposée donne $\pi = 3,1464466094\dots$ ce qui approche π mais est contraire à la valeur admise pour π dès la quatrième décimale. Bien sûr, la preuve

Une méthode plus efficace mais heureusement assez peu mise en œuvre jusqu'à présent vient d'être identifiée par Guillaume Cabanac, professeur à l'université de Toulouse, et Lonni Besançon, de l'université de Linköping, en Suède, et quelques autres chasseurs de fraudes scientifiques. Guillaume Cabanac étudiait un article contenant des expressions *torturées*, c'est-à-dire peu naturelles et faisant soupçonner un plagiaire qui modifie le texte copié pour l'éloigner du texte volé. Il a alors remarqué des

L'inflation des publications

Le nombre d'articles scientifiques publiés chaque année (ici, la production des principaux éditeurs scientifiques) ne cesse de croître et augmente plus rapidement que le nombre de chercheurs. Ce phénomène d'inflation se produit même en excluant les revues prédatrices qui, de leur côté, ont vu leur nombre exploser au point qu'on en compte aujourd'hui plusieurs milliers. Entre 2016 et 2022 un calcul soigneux permet d'évaluer à 47 % l'augmentation du nombre d'articles scientifiques des meilleures revues.



n'est pas correcte. Dans le cours de la démonstration, Reddy admet subrepticement une égalité entre la longueur du périmètre d'un cercle et le tracé d'une partie du périmètre d'un carré, ce qui revient à admettre l'expression qu'il donne pour π , cette égalité entre longueurs n'étant nullement prouvée; en clair, il prouve sa prétendue formule en l'admettant au détour d'une phrase.

Références furtives

Il existe cependant des méthodes plus élaborées que le recours aux revues prédatrices pour étendre son influence. Pour faire croître les citations aux articles de l'auteur E, la méthode de base consiste bien sûr pour E à se citer lui-même. Une variante de l'indice h ne prenant pas en compte les autocitations évite ce biais. Plus subtil, E peut convenir avec F de citer ses papiers même sans raison, avec en retour la promesse que F citera les articles de E. Implicites ou explicites, ces réseaux de soutien mutuel qui concernent souvent bien plus de deux auteurs sont en réalité impossibles à distinguer automatiquement des pratiques de citations croisées et justifiées entre scientifiques et équipes travaillant sur les mêmes sujets.

statistiques bizarres: un article était cité 107 fois, alors qu'il n'avait été téléchargé que 62 fois. Le rapport habituel est totalement différent, car un article est en général téléchargé 10 ou même 50 fois plus qu'il n'est cité. En tentant de comprendre, les chasseurs de fraude ont découvert une astuce insoupçonnée: le trucage des métadonnées d'un article. Partant du texte d'une publication, par exemple sous forme de fichier pdf, certains opérateurs centralisant les informations sur les publications scientifiques y associent des métadonnées reprenant des éléments caractéristiques de l'article, dont en particulier son auteur et la bibliographie de l'article. Plutôt que de calculer le nombre de références faites à un article à partir du fichier même de l'article, il est plus simple de travailler à partir de ces bases de métadonnées pour par exemple calculer l'indice h d'un chercheur ou d'une chercheuse. La manipulation découverte est que les métadonnées de nombreux articles ont été truquées et les citations qu'elles indiquent ne correspondent pas aux citations véritablement faites dans le texte des articles. L'article bizarre en réalité n'avait jamais été cité par personne, seul son ajout frauduleux dans la bibliographie des métadonnées d'autres articles permettait d'arriver

aux 107 citations, en réalité purement fictives. La fraude était invisible, car personne ne va vérifier que les métadonnées d'un article sont correctement générées. L'origine de cette manipulation, nommée maintenant «la méthode des références furtives», n'est pas très facile à identifier précisément, mais des recherches systématiques ont été entreprises. Dorénavant la méfiance concernant les métadonnées sera nécessaire.

Cette tricherie rappelle une farce opérée par Cyril Labbé, professeur à l'université de Grenoble, dans le but de montrer que si une base de données d'articles est trop tolérante, il est alors facile de la manipuler.

Production automatique

En créant l'auteur imaginaire Ike Antkare – à prononcer *I can't care*, qui signifie «Je m'en moque» – et en faisant engendrer 100 articles par un programme produisant de faux articles scientifiques: il en existe et bien sûr l'IA générative rend aujourd'hui plus facile encore la production automatique de faux articles. Celui qui a été utilisé se nomme SciGen. Cyril Labbé a réussi à berner la base de données bibliographiques de Google Scholar. Pour cela Ike Antkare était l'auteur de chacun des 100 faux articles. Ces articles en citaient de déjà connus de Google Scholar, ce qui est un critère pour qu'ils soient considérés comme acceptables par cette base, et bien sûr les 100 articles se citaient mutuellement tous. En créant une page internet permettant le téléchargement de tous les articles d'Ike Antkare, le système de fouille d'internet de Google a pris connaissance des 100 articles. Pour un système de calcul de l'indice h utilisant Google Scholar cela aurait dû conduire à attribuer l'indice h de 100 au prétendu auteur. Sans qu'on sache bien pourquoi, ce n'est pas tout à fait ce qui s'est passé et le système de calcul Scholarometer, basé sur Google Scholar, ne lui a attribué qu'un indice h de 94. Quand la farce a été montée, un tel indice h classait Ike Antkare en 21^e position de tous les scientifiques derrière Sigmund Freud, le premier (indice h de 183), mais devant Albert Einstein (indice h de 84). En utilisant 200 articles factices plutôt que 100, Ike Antkare aurait été le numéro un de l'indice h pour la base de Google Scholar!

Notons cependant que ce remarquable et amusant exploit n'aurait pas pu aboutir avec d'autres bases de données comme celles de Scopus ou Web of Science, qui ne prennent en considération que les articles parus dans des revues en nombre limité, en excluant en particulier les articles des revues prédatrices et les articles simplement déposés sur des pages internet.

Le monde scientifique avec ses multiples publications est extrêmement délicat à suivre.

Toute méthode automatique de mesure de ses acteurs doit être contrôlée, validée et examinée avec méfiance; les histoires d'Ike Antkare, des références furtives et bien d'autres l'ont clairement démontré. Certaines personnes et certaines publications échapperont cependant toujours aux meilleurs protocoles quel que soit le soin mis à les concevoir. Voici quelques exemples de cas singuliers ou extrêmes montrant les difficultés peut-être insurmontables que rencontrera toute entreprise de notation algorithmique.

— Revues prédatrices et sites de «preprints» —

Les revues qualifiées de «prédatrices» ne disposent pas de système véritable d'évaluation et de rejet par des experts du domaine concerné. Elles acceptent de tout publier pourvu que les frais demandés soient acquittés. Bien sûr, de telles revues sans contrôle sont considérées avec hostilité et mépris par la communauté scientifique, qui n'a cependant pas réussi à éviter leur multiplication depuis vingt ans.

À côté de ces revues, d'autres dispositifs de publication sans révision par les pairs jouent un rôle très important aujourd'hui: ce sont les sites dits de *preprints*, rendant accessibles les prépublications d'articles produits pour être ultérieurement soumis à révision.

Avant que la revue à laquelle un article est soumis prenne la décision de le publier ou de le rejeter – ce qui peut demander de longs mois et même des années –, les chercheurs et chercheuses souhaitent rendre accessibles les résultats de leur travail grâce à ces publications non encore expertisées, ce qui est utile pour tout le monde puisque cela accélère la circulation de l'information.

Le site arXiv est le plus important de cette catégorie. Hébergé à sa création au Los Alamos National Laboratory, il a été transféré à l'université privée Cornell, dans l'État de New York, aux États-Unis, qui le gère aujourd'hui.

Publier un article seul ou à plusieurs ne signifie pas la même chose surtout si les signataires sont nombreux. Les 5 154 auteurs (oui 5 154!) de l'article proposant une valeur pour la masse du boson de Higgs publié en 2015 dans la revue *Physical Review Letters* ne peuvent certainement pas être considérés tous de la même façon; ils n'ont probablement pas tous participé à l'écriture d'un seul mot de l'article qui ne faisait que 33 pages (dont les noms et affiliations des auteurs occupaient plus de la moitié). Quel sens peut bien avoir de mettre cet article dans sa liste de publications?

L'usage accepté dans certaines disciplines que le directeur cosigne tous les articles des membres de son laboratoire conduit aussi à des absurdités comme celle qui concerne Didier Raoult, le célèbre défenseur de l'hydroxychloroquine pour soigner le Covid-19. Il est l'auteur de plus de 3 000 articles,

Le facteur d'impact des revues

Le facteur d'impact FI (*impact factor* en anglais) est un indicateur numérique lié à la visibilité d'une revue scientifique. Le FI pour l'année A d'une revue est égal à la moyenne du nombre de citations des articles de cette revue publiés durant les deux années précédentes. Il dépend bien sûr de la base de données d'articles utilisée pour compter les références, mais c'est le FI basé sur celle du Web of Science et publié chaque année par la *Journal of Citation Reports* qui est presque toujours mentionné.

Les deux plus prestigieuses revues scientifiques, *Nature* et *Science*, ont des facteurs d'impact variant autour de 40. Seules 2% des revues ont un FI dépassant 10. Un quart des revues a un FI inférieur à 1. La durée de deux ans dans la prise en considération des articles dont on compte les références est souvent jugée trop courte, car une recherche exige

parfois du temps avant d'être reconnue et donc citée. Un autre défaut du facteur d'impact si on l'utilise pour évaluer l'importance d'un travail de recherche est qu'un article très médiocre peut avoir été publié dans une revue avec un bon facteur d'impact puisque celui-ci est fondé sur une moyenne. Inversement, un article très important (comme celui de Satoshi Nakamoto sur le bitcoin) serait considéré comme mauvais puisqu'il n'a été publié dans aucune revue ayant un facteur d'impact.

L'évaluation des scientifiques repose de plus en plus sur des indicateurs bibliométriques comme l'indice *h* et le FI, qui mesurent à la fois leur productivité et l'impact de leurs publications *via* les citations.

soit plus de 5 par mois durant toute sa carrière... et le plus souvent dans des revues dont il était membre du comité de rédaction.

Indice modeste pour Perelman

Les trois articles du mathématicien russe Grigori Perelman démontrant la conjecture de Poincaré n'ont jamais été publiés dans une revue reconnue avec comité de lecture, mais seulement déposés sur le site arXiv de prépublications. Cette démonstration reconnue aujourd'hui correcte par les spécialistes a valu à son auteur une médaille Fields (un équivalent du prix Nobel, pour les mathématiques) et un prix Clay de 1 million de dollars... qu'il a tous les deux refusés! Malgré son exploit exceptionnel les indices *h* et autres de Perelman ne sont pas très bons quand ils sont calculés rigoureusement en se fondant sur des bases de données strictes. Google Scholar, pourtant très généreux, ne lui donne qu'un indice *h* de 8, c'est-à-dire moins qu'un jeune chercheur moyen.

Un autre cas remarquable est celui de Satoshi Nakamoto. Derrière ce pseudonyme se cache vraisemblablement une équipe de plusieurs spécialistes de monnaies numériques. Dans un article de 2008, Satoshi Nakamoto a décrit le principe de fonctionnement de la monnaie cryptographique bitcoin. Il a aussi

contribué à sa mise en fonctionnement. L'article qu'on peut considérer d'une grande importance puisque la capitalisation de Bitcoin aujourd'hui dépasse mille milliards de dollars, et que le modèle de fonctionnement du bitcoin a été copié des centaines de fois, n'a jamais été publié dans une revue avec comité de programme. Aux yeux de nombreux systèmes de calcul des indices *h*, *g*, *e*, etc., il ne vaut donc rien.

L'anonymat de Nakamoto est un problème pour ces indices, mais il concerne aussi d'une façon différente le mathématicien Nicolas Bourbaki. On sait que derrière ce pseudonyme se cachent (plus ou moins!) des mathématiciens dont la liste évolue dans le temps. Que l'indice *h* de Bourbaki n'ait aucun sens est évident puisque derrière ce nom il y a une collectivité sans cesse renouvelée. Mais le problème concerne tous les mathématiciens qui participent à l'écriture de ces travaux: ils ne sont crédités d'aucun des résultats exceptionnels, universellement estimés et très largement cités qu'ils ont contribué à élaborer. Ils sont donc pénalisés si on calcule leurs indices *h*, *g*, *e*, etc.

Plus généralement, certaines recherches de nature militaire ou donnant lieu à des secrets industriels ne sont tout simplement pas publiées, quelle que soit l'importance des résultats auxquels elles sont parvenues, ce qui bien sûr fausse totalement les évaluations menées avec les systèmes automatiques des scientifiques y ayant participé.

Évaluer les chercheurs scientifiques mais aussi noter les revues et congrès pour déterminer les meilleurs et ceux auxquels il faut accorder le plus d'importance et de confiance, voilà des tâches probablement impossibles dans l'absolu et où l'utilisation de méthodes automatiques doit toujours s'accompagner d'un regard vigilant et méfiant à l'extrême. ■

BIBLIOGRAPHIE

- A. Abalkina et al.**, « Stamp out paper mills » — science sleuths on how to fight fake research, *Nature*, 2025.
- L. Besançon et al.**, Detection of metadata manipulations: Finding sneaked references in the scholarly literature, *arXiv* (preprint), 2025.
- M. Hanson et al.**, The strain on scientific publishing, *Quantitative Science Studies*, 2024.
- L. Besançon et al.**, Sneaked references: Fabricated reference metadata distort citation counts, *Journal of the Association for Information Science and Technology*, 2024.
- G. Cabanac et al.**, The « Problematic Paper Screener » automatically selects suspect publications for post-publication (re)assessment, 7th WCRI, 2022.
- Y. Gingras**, The transformation of the scientific paper: From knowledge to accounting unit, dans M. Biagioli et A. Lippman (éd.), *Gaming the Metrics*, MIT Press, 2020.
- N. Chevassus-au-Louis**, *Malscience. De la fraude dans les labos*, Le Seuil, 2016.
- www.predatoryjournals.org/thelist
Liste de revues prédatrices