

Accueil > Blog > Mesures et démesure de la publication scientifique

Mesures et démesure de la publication scientifique



ACTUALITÉS DU COMITÉ 21/10/2024

Benoît Pier (Directeur de recherche au CNRS au Laboratoire de mécanique des fluides et d'acoustique – LMFA) et Laurent Romary (Directeur de recherche, Inria) nous proposent un billet sur l'augmentation récente du nombre de publications scientifiques et son effet sur la recherche scientifique, à partir de l'analyse faite dans le "preprint" rendu public sur arXiv : "Hanson, M. A., Barreiro, P. G., Crosetto, P., & Brockington, D. (2023). The strain on scientific publishing. arXiv preprint arXiv:2309.15884. <https://doi.org/10.48550/arXiv.2309.15884>"

Une étude publiée dans arXiv, intitulée "*The strain on scientific publishing*" revient sur

l'explosion du nombre de publications scientifiques et cherche à la documenter précisément avec des données quantitatives. L'objectif est d'identifier les ingrédients de cette croissance et d'en appréhender les conséquences sur l'activité de recherche, notamment le travail de validation par les pairs et la veille scientifique.

Les auteurs sont des spécialistes de bibliométrie et se basent sur un ensemble de sources pour effectuer leurs analyses quantitatives : surtout Scopus et Web of Science, mais aussi moissonnage direct depuis les pages web de différentes revues et maisons d'édition. La période analysée couvre les dix dernières années, avec un focus particulier sur la période 2016-2022.

Le premier résultat concerne le nombre total d'articles publiés, qui est très proche d'une croissance exponentielle (avec un taux +5.6% par an). Même en prenant en compte l'augmentation du nombre de chercheurs sur cette période, on en déduit que le temps consacré à l'obtention des résultats, leur validation et leur vérification par les pairs a significativement diminué.

Les auteurs font état de plusieurs facteurs pouvant expliquer cette croissance, incluant notamment une plus grande accessibilité aux publications scientifiques dans les pays du «Sud global ». Cependant, certains facteurs semblent prédominants, avec en tête les nouvelles pratiques éditoriales. En ventilant le nombre d'articles publiés annuellement par maison d'édition, les auteurs démontrent que cette forte augmentation est essentiellement due à quelques éditeurs qui ont énormément

accru leur volume de publications : de manière flagrante, MDPI et Frontiers sont sur des trajectoires très différentes de la plupart des autres éditeurs. En calculant le nombre annuel moyen d'articles publiés par revue scientifique, les auteurs font le constat que la croissance (significative, mais pas vertigineuse) des éditeurs classiques est surtout due à l'élargissement de leur catalogue de revues, alors que la très forte croissance de Frontiers et MDPI résulte d'une augmentation explosive du nombre d'articles par revue. Notons que ces deux éditeurs, apparus plus récemment, se développent grâce aux frais de publication payés par les auteurs.

Pour creuser cette question, les auteurs ont également comparé le nombre d'articles publiés dans des numéros spéciaux avec ceux publiés dans les numéros réguliers. Le constat est fait que pour MDPI, Hindawi et Frontiers, la très forte croissance est presque exclusivement due à l'augmentation des articles publiés dans des numéros spéciaux : notamment en 2022, la très grande majorité des articles de ces éditeurs a été publiée dans des numéros spéciaux.

Pour caractériser plus finement les pratiques de publication, les auteurs ont effectué une analyse statistique de la durée du cycle de relecture par les pairs, c'est-à-dire du temps écoulé entre la soumission et l'acceptation. Ces données sont difficilement disponibles, mais les auteurs ont pu les moissonner pour un nombre suffisant d'articles pour en déduire les distributions statistiques par grands éditeurs, et comparer leur évolution

sur les années 2016, 2019 et 2022. Cette analyse statistique montre qu'en général ces distributions n'ont guère évolué entre 2016 et 2022 et sont asymétriques, avec une longue traîne aux temps longs, correspondant aux « événements rares » que sont les articles particulièrement difficiles à faire expertiser. Mais là encore, Frontiers, MDPI et Hindawi ont un comportement inhabituel : en 2022, la distribution des temps de relecture se resserre autour des temps nettement plus courts qu'en 2016 et s'accompagne de la perte de la longue traîne. Cette triade de maisons d'édition est donc passée à un processus beaucoup plus rapide et homogène, alors que les disciplines couvertes sont très variées et la qualité des manuscrits probablement tout aussi variable que par le passé.

Ainsi, cette étude met en avant deux facteurs principaux qui alimentent l'explosion du nombre d'articles publiés :

- Les éditeurs plus anciens accroissent surtout leur catalogue, avec quand même aussi un accroissement du nombre d'articles par revue.
- La triade (MDPI, Frontiers, Hindawi) utilise majoritairement des numéros spéciaux pour augmenter très fortement la quantité d'articles sortis, et ce phénomène s'accompagne d'un raccourcissement significatif du temps entre soumission et acceptation.

Cette étude n'a pas manqué de faire réagir, notamment Frontiers par un billet de blog qui conteste les chiffres, accuse les auteurs d'avoir des préjugés, et continue par ailleurs

de faire l'éloge de la croissance exponentielle. Les auteurs de la présente étude ont fait remarquer que, contrairement à d'autres éditeurs, Frontiers n'a pas fourni les données demandées, et ils démontent méthodiquement les erreurs d'analyse de Frontiers dans leur billet de blog.



ce genre d'études, seront les seules qui

garantissent la validité scientifique des

résultats obtenus. Nous pouvons **RESSOURCES**

ACTUALITES

EN

regretter que ce travail s'appuie en grande partie sur les bases de données fermées de

Scopus et du Web of Science. Néanmoins,

nous avons tenté l'exercice de reproduire

QUI SOMMES NOUS ?

certaines aspects en interrogeant les données

ouvertes de la base OpenAlex et nous

retrouvons bien exactement les mêmes

tendances en ce qui concerne l'évolution du

nombre d'articles publiés par an, par journal

et par éditeur (voir figure ci-dessous). Mais il

n'est actuellement pas possible de refaire

toute cette étude fouillée avec des données

ouvertes.

Espérons que la déclaration de Barcelone sur

l'ouverture des informations sur la recherche

(dont le Comité pour la science ouverte est

signataire) puisse faciliter cela dans le futur.

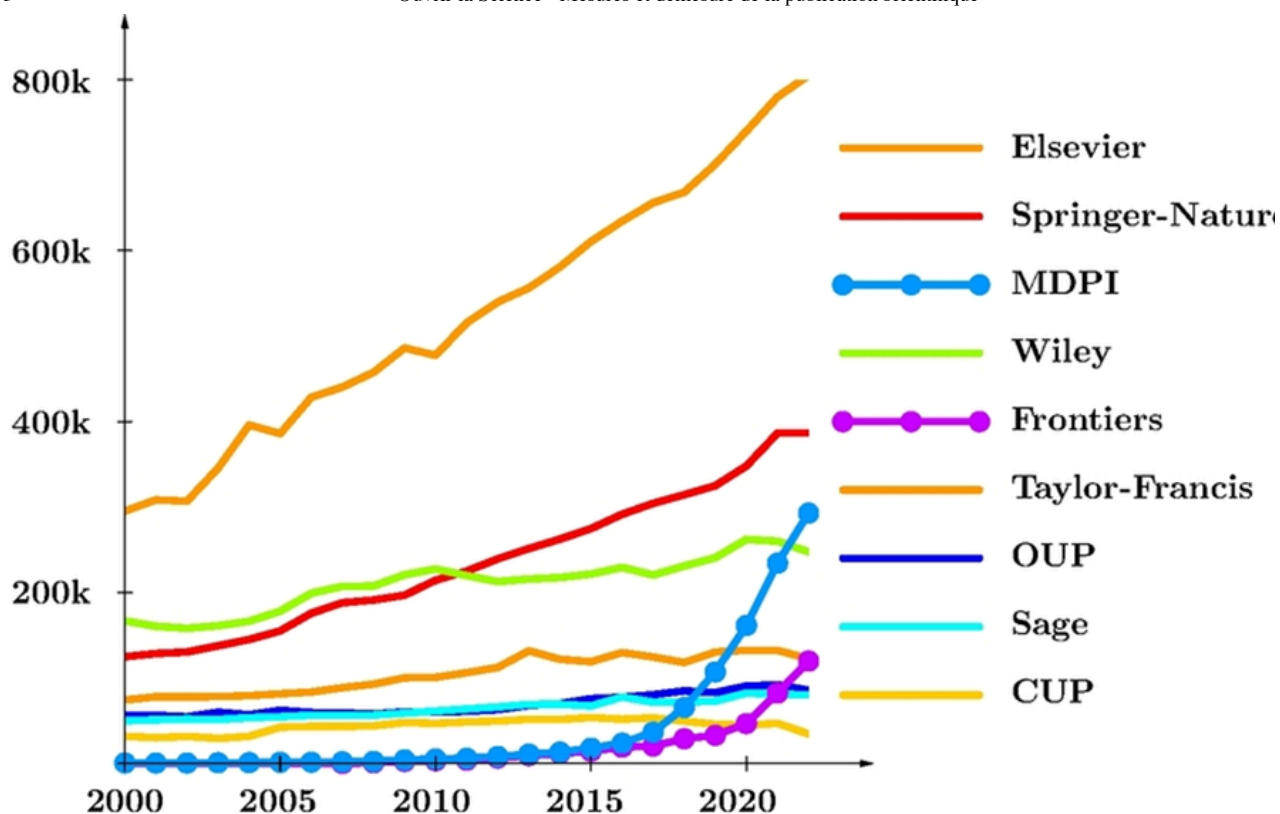


Figure 1: Évolution du nombre total d'articles publiés par an et par éditeur, données issues de OpenAlex.

En conclusion, sur la base des résultats de cette étude, nous pouvons nous poser la question de savoir si le système actuel est soutenable. La réponse est clairement, non. Une croissance exponentielle du nombre de publications n'est pas compatible avec le maintien de la qualité scientifique et la confiance dans les résultats, garantis par le minutieux travail de la relecture par les pairs.

D'autant plus que les effectifs de la communauté scientifique ne sont pas ou peu en augmentation, si l'on se réfère à l'évolution du nombre de doctorantes et de doctorants diplômés dans les pays de l'OCDE, qui est en légère récession depuis 2018 comme l'observent les auteurs de l'étude présentée ici. Or, pour maintenir un haut niveau d'exigence, il faut que chaque article publié

soit validé par au moins deux relecteurs indépendants qui puissent y consacrer suffisamment de temps ; et les comités éditoriaux de la plupart des revues scientifiques rencontrent de plus en plus de difficultés pour trouver des relecteurs. En prenant en compte le taux de rejet, nous proposons d'énoncer la règle d'or suivante : pour entretenir un régime de publication de qualité, il faut que chaque scientifique expertise environ dix fois plus d'articles de ses pairs qu'il ou elle n'en écrit.

Devant l'actuelle inflation de la publication scientifique, certains préconisent de développer de nouveaux outils (par exemple avec de l'« intelligence artificielle »), pour faire le tri dans ce qui est pertinent, évaluer ce qui en vaut la peine, identifier les textes eux-mêmes générés par des machines, etc. Mais tout comme pour le dérèglement climatique, le dérèglement de l'édition scientifique a peu de chances d'être résolu par des moyens purement techniques, si on ne s'appuie pas d'abord sur l'intelligence humaine, aidée par un peu de sobriété.

En même temps, il faut que les établissements, universités et organismes de recherche, ainsi que les agences de financement soutiennent sans ambiguïté une réforme de l'évaluation basée sur les principes de DORA et COARA (voir le chapitre français de COARA : <https://coara.fr/>), en bannissant des indicateurs comme le « h-index » et le « journal impact factor » et en encourageant les CV narratifs. L'évolution des pratiques de publication vers un système plus

vertueux ne peut se faire qu'avec des pratiques d'évaluation cohérentes.

Plutôt que d'entretenir l'actuelle course à la publication, il vaut mieux consacrer notre temps à faire de la meilleure science et donc publier moins et mieux.

Mise à jour du 13 novembre 2024.

Depuis la publication de ce billet, la prépublication « The strain on scientific publishing » est devenue une publication et *Le Monde* a traité le sujet. Nous indiquons les références ci-dessous.

- Hanson, Mark A., Pablo Gómez Barreiro, Paolo Crosetto, et Dan Brockington. « The strain on scientific publishing ». *Quantitative Science Studies*, 8 novembre 2024, 1-21.
<https://doi.org/10.48550/arXiv.2309.15884>
- « L'inflation du nombre de publications scientifiques interroge ». 12 novembre 2024 – accès réservé aux abonnés.
https://www.lemonde.fr/sciences/article/2024/11/12/l-inflation-du-nombre-de-publications-scientifiques-interroge_6389778_1650684.html

