

ScreenIT : quand l'automatisation du peer review se met au service de la science ouverte

 openscience.pasteur.fr/2026/04/27/screenit-quand-lautomatisation-du-peer-review-se-met-au-service-de-la-science-ouverte/

CeRIS - Institut Pasteur

27 avril 2026



[Une postdoc vous raconte](#)

[Version anglaise accessible ici : « [ScreenIT: harnessing automated peer-review to support Open Science](#) »]

Connaissez-vous la communauté de [ScreenIT](#) (anciennement l'*Automated Screening Working Group*) ? Depuis 2020, ce collectif de chercheur·se·s et d'informaticien·ne·s développe des outils visant à améliorer la rigueur et la reproductibilité de la recherche scientifique. Ce travail a débuté face à l'avalanche de papiers prépubliés sur bioRxiv et medRxiv concernant la pandémie de COVID-19. Ces données constituaient une ressource cruciale pour la communauté scientifique et médicale, mais leur volume et leur rythme de production étaient tels que le flux de préprints dépassait les capacités du *peer review* (entraînant un délai important avant validation). ScreenIT a donc développé un pipeline [pour analyser automatiquement ces préprints](#) et identifier les problèmes communs qu'ils présentaient. Les rapports d'analyse des préprints étaient ensuite partagés [sur hypotheses.is](#). Parmi leurs critères ? La présence de code et de données accessibles.

Faisons un bond en avant jusqu'en 2026. Heureusement, la pandémie est terminée, mais le problème de la surproduction de publications par rapport à nos capacités de *peer review* n'a fait qu'empirer avec les avancées de l'IA. Par exemple, le groupe Wiley a enregistré une [hausse de 25% des soumissions à ses journaux](#) au premier trimestre de 2025. En octobre dernier, [une enquête de Cambridge University Press](#) a révélé que 81%

des 3 000 chercheur·se·s interrogé·e·s considéraient que « l'augmentation du nombre de publications scientifiques avait mis le système de *peer review* sous pression ». Par ailleurs, on estime que [10% des publications de 1999 à 2024 dans le domaine du cancer sont des fraudes](#) provenant de [paper mills](#). Pour identifier les articles frauduleux, [un outil récemment publié](#) dans *BMJ* détecte automatiquement certains motifs typiques (mauvaises descriptions de réactifs, images manipulées ou recyclées, similitudes au niveau du texte et de la mise en page, etc.). Mais pour analyser la crédibilité et la reproductibilité des articles légitimes, un outil capable d'analyser la transparence du travail et des données semble plus pertinent.

En février, les membres de ScreenIT ont publié [une étude](#) portant sur leur pipeline actuel, dans laquelle iels comparent 11 outils à l'aune de 9 critères de transparence, en les appliquant à 1 500 manuscrits en libre accès de PubMed Central. Avec, bien sûr, des évaluateur·rice·s humain·e·s pour établir un *gold standard* de chaque tâche. Les auteur·rice·s cherchaient à comprendre non seulement quel outil était le meilleur pour chaque tâche, mais aussi si une combinaison d'outils pouvait être plus performante. Les résultats sont mitigés : pour certaines tâches, comme l'évaluation de la transparence sur les critères d'inclusion des participants dans les études et essais cliniques, la combinaison de plusieurs outils complémentaires donnait de meilleurs résultats. Pour d'autres, comme l'identification des logiciels utilisés, un outil, [SoftCite](#), est sorti gagnant. Personnellement, le critère de transparence qui m'intéresse le plus est celui concernant l'accessibilité au code. Contre toute attente, l'outil qui s'est révélé le plus efficace pour cette tâche est [ODDPub](#), plus performant que l'alternative commerciale IA [SciScore](#).

Nous sentons toujours l'influence du contexte pandémique dans la collection d'outils de ScreenIT : beaucoup d'entre eux portent sur les particularités des études et essais cliniques, ce qui limite leur application à d'autres types d'articles scientifiques. Néanmoins, ScreenIT réalise également des analyses de bonnes pratiques plus générales, comme l'utilisation de la *randomisation* ou d'une méthode en simple ou double aveugle, ou l'évaluation de la puissance de l'étude en fonction de la taille de l'échantillon. Malgré l'intérêt de ces outils, je n'ai pas trouvé d'accès au pipeline global. Pour une utilisation personnelle, il semble donc que nous devions chercher chaque outil individuellement.

Et à l'échelle d'une communauté ou d'une entreprise, ces outils seront-ils utilisés ? Pour un journal, la réponse est déjà oui : Jessica Leight, chercheuse à l'*International Food Policy Research Institute* aux Etats-Unis et éditrice académique chez PLOS One, confirme l'adoption de ScreenIT par le journal dans [un post LinkedIn](#). Ces rapports ScreenIT sont déjà disponibles dans certains *peer review* en libre accès de PLOS One, avec un logo représentant un petit robot (par exemple [ce rapport](#) associé à [cet article](#)). Reste à voir comment cet ajout dans le processus de *review* aidera les éditeur·rice·s et auteur·rice·s dans les démarches de publication.

Prenons un peu de recul. Hors du contexte de PLOS One, je pense que ScreenIT pourrait parfaitement remplir son rôle original : trier rapidement les préprints et identifier ceux dont les méthodes sont suffisamment solides et transparentes pour être considérées, avant

même le *peer review*. Je pourrais donc l'imaginer intégré aux plateformes bioRxiv ou medRxiv. Ainsi ScreenIT permettrait de vérifier rapidement une série de critères de transparence, mais je ne pense pas qu'il pourrait se substituer aux pair·e·s humain·e·s. Et pourquoi pas utiliser ScreenIT pour évaluer la transparence des travaux d'un·e scientifique ? Cela pourrait constituer une première étape pour [intégrer la reproductibilité](#) dans les décisions en matière de recrutement, de promotion ou de financement.

[Caitlin Martin](#), postdoctorante à l'Institut Pasteur

Références :

- Weissgerber et al. 2021. Automated screening of COVID-19 preprints: can we help authors to improve transparency and reproducibility? Nature Medicine. 27:6-7. <https://doi.org/10.1038/s41591-020-01203-7>
- Scancar et al. 2026. Machine learning based screening of potential paper mill publications in cancer research: methodological and cross sectional study. BMJ. 392. <https://doi.org/10.1136/bmj-2025-087581>
- Eckmann et al. 2026. Use as directed? A comparison of software tools intended to check rigor and transparency of published work. PLOS One. <https://doi.org/10.1371/journal.pone.0342225>