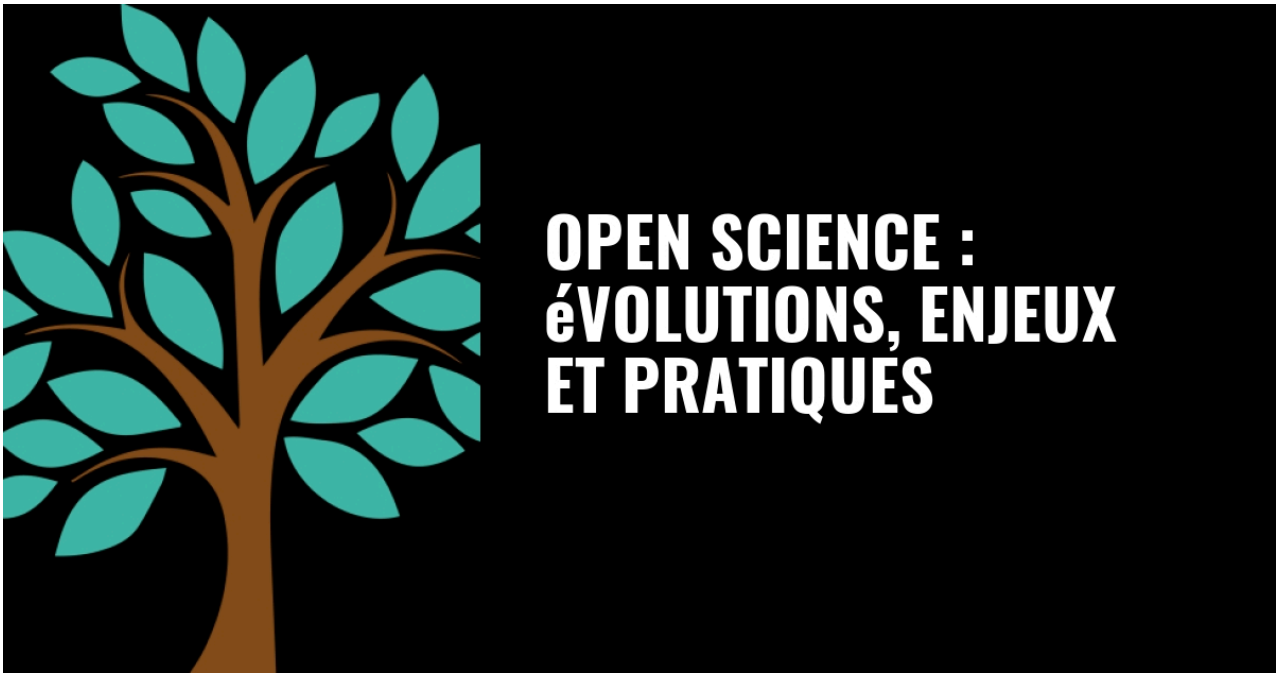
Et si la reproductibilité de nos données faisait partie de notre évaluation ? – Open science : évolutions, enjeux et pratiques

openscience.pasteur.fr/2026/03/30/et-si-la-reproductibilite-de-nos-donnees-faisait-partie-de-notre-evaluation/

CeRIS - Institut Pasteur 30 mars 2026



[Version anglaise accessible ici : « What if the reproducibility of our science was part of our evaluation? »]

It's that time of year again – c'est le moment de l'année où les jeunes chercheur·se·s attendent avec impatience (et parfois avec crainte) les résultats des concours CNRS, INRAE, et INSERM. Nous constatons tous·tes l'importance de nos publications dans ces évaluations, mais existe-il aussi une place pour notre engagement dans la science ouverte ?

Face à la « crise de la reproductibilité scientifique » souvent évoquée comme conséquence de la pression à publier beaucoup et vite, Sarah Cohen-Boulakia, professeure et bio-informaticienne à l'université Paris-Saclay, suggère d'inclure la reproductibilité de nos résultats comme critère d'évaluation, surtout dans les domaines bio-informatiques. Ayant déjà vécu la frustration de ne pas pouvoir utiliser des outils ou des données publiés (protocoles incompréhensibles et sans annotations, dépendances d'outils obsolètes, liens morts, etc.) même *avant* d'essayer de reproduire leurs résultats, j'aime beaucoup cette idée.

Et je ne suis pas la seule. Une récente enquête a ainsi montré que 90% des chercheur·se·s évaluateur·rice·s considèrent la « crédibilité » des travaux comme un critère d'évaluation « très important ». Cette étude a été menée auprès de 485 chercheur·se·s en biologie ayant fait partie de comités d'évaluation (de financement,

recrutement et promotion) au cours des deux dernières années. Par recherche « crédible », on entend une recherche fiable, bien conçue, transparente, éthique, et avec des méthodes exhaustives. À titre de comparaison, seul·e·s 54% des chercheur·se·s interrogé·e·s considèrent que « l'impact de la recherche » est important. Néanmoins, si 45% se disent satisfait·e·s de leur capacité à évaluer cet impact, seul·e·s 38% le sont de leur capacité à évaluer la crédibilité. En conséquence, pour évaluer la crédibilité, 57% des chercheur·se·s se basent sur la réputation de la personne et de son laboratoire et/ou sur le facteur d'impact des journaux dans lesquelles elle publie !

Pour avoir un premier baromètre de la transparence et, par extension, de la « crédibilité » du travail, le niveau de partage des données, des codes, et des méthodes semble être un point de départ naturel. Cependant, évaluer dans quelle mesure les publications scientifiques d'un·e candidat·e sont alignées avec les standards de la science ouverte s'avère être une tâche particulièrement compliquée. Dans l'enquête mentionnée précédemment, seul·e·s 30% des chercheur·se·s se disaient très satisfait·e·s de leur capacité à évaluer la transparence des publications des candidat·e·s. La faible considération pour la science ouverte dans le recrutement actuel pourrait-elle être liée à cette difficulté d'évaluation ?

Comment donc évaluer notre propre reproductibilité ? Ou plus simplement, comment évaluer le niveau de transparence d'un papier ?

Je n'ai pas encore de réponse satisfaisante à cette question, mais nous commençons à avoir des pistes. PLOS a développé ses propres mesures, les OSI (*Open Science Indicators*) basées sur six critères en accord avec les principes FAIR (*Findable, Accessible, Interoperable, Reusable)*. Cependant, ces métriques restent à l'échelle des journaux et ne permettent pas d'évaluer des articles individuels. À l'échelle des articles, PLOS a créé en 2022 un badge expérimental « *Accessible Data* », qui s'est élargi en 2023. Néanmoins, ce badge est attribué à la seule condition qu'un lien vers un entrepôt soit inclus dans le *Data Availability Statement* de l'article, ce qui ne permet pas de savoir si la totalité ou une partie seulement des données est accessible, ni dans quelles conditions. Les articles pourvus d'un tel badge peuvent donc être très hétérogènes en termes d'ouverture. Allant une étape plus loin, l'ACM (*Association for Computing Machinery*) a développé des badges qui indiquent d'abord si les « *artefacts* » (données, codes, logiciels, etc.) de l'article sont librement accessibles, et ensuite, si les résultats ont été reproduits et/ou répliqués ailleurs (nous vous expliquons plus en détail les badges Science Ouverte ici). Mais ces badges concernent uniquement les journaux de l'ACM, et ne nous aideront pas pour évaluer le profil de « crédibilité » global d'un chercheur.

Alors que faire ?

Il semblerait que ce soit la réponse qui revient sans cesse en ce moment : il existe de nouveaux outils d'IA pour cette tâche. Plusieurs outils de ScreenIT, par exemple, évaluent l'aspect « *Open Science »* de nos articles. Cette fois, je ne les ai pas testés pour vous, car une équipe a déjà effectué le travail !

Stay tuned, je le traiterai dans mon prochain article ! Nous verrons dans quelle mesure ces outils pourront nous aider à révolutionner nos évaluations…

Caitlin Martin, postdoctorante à l'Institut Pasteur

Références :

- Sarah Cohen-Boulakia : « On publie trop et trop vite ». Interview par Lucile Veissier, TheMetaNews, 13 février 2026
- Hrynaszkiewicz et al. 2026. A survey of how biology researchers assess credibility when serving on grant and hiring committees. PeerJ 14:e20502 https://doi.org/10.7717/peerj.20502
- Khan et al. 2022. Open science failed to penetrate academic hiring practices: a cross-sectional study. Journal of Clinical Epidemiology 144:136-143 https://doi.org/10.1016/j.jclinepi.2021.12.003

Tagué Evaluation de la recherche, Reproductibilité, Transparence