
Pollution de la littérature scientifique : détection participative d'expressions torturées révélatrices d'articles frauduleux

Guillaume Cabanac

IRIT UMR 5505 CNRS, Université Paul Sabatier, Toulouse, France
guillaume.cabanac@irit.fr

RESUME.

Nous avons découvert des milliers de publications non fiables dans les catalogues des maisons d'édition de premier plan : Elsevier, Springer et Wiley, notamment. Publiés et souvent vendus, ces pseudo-articles générés par ordinateur ou assemblés par des paper mills¹²³ tels des patchworks sont trahis par la présence d'« expressions torturées » dénuées de sens. Cet exposé présentera la plateforme 'Problematic Paper Screener'⁴ pour identifier cette pollution affectant la littérature scientifique. Avec d'autres détectives scientifiques et lanceurs d'alerte, nous l'employons pour ré-évaluer les 6 000 articles identifiés à ce jour et les signaler sur la plateforme d'évaluation post publication PubPeer pour les faire rétracter. Cette initiative bénévole de fact-checking participatif détecte de nouvelles phrases torturées et méconduites qui sont intégrées au système développé, conduisant à étendre le détecteur par effet boule de neige. Les 6 000 articles problématiques sont parus majoritairement depuis 2014 et font l'objet de 40 000 citations en tout. Des centaines d'entre eux sont abusivement cités, sans logique apparente, indice d'une manipulation visant à augmenter le nombre de citations de certains fraudeurs.

Cet exposé reprendra notre article paru dans le Bulletin of the Atomic Scientists créé en 1945 par les scientifiques du projet Manhattan, traitant de « la sécurité mondiale et les questions de politique publique, en particulier celles liées aux dangers posés par les armes nucléaires et autres armes de destruction massive » (Wikipedia⁵).

¹ <https://forbetterscience.com/2020/01/24/the-full-service-paper-mill-and-its-chinese-customers/>

² <https://www.nature.com/articles/d41586-021-00733-5>

³ <https://forbetterscience.com/2021/05/26/the-chinese-paper-mill-industry-interview-with-smut-clyde-and-tiger-bb8/>

⁴ <https://www.irit.fr/~Guillaume.Cabanac/problematic-paper-screener>

⁵ https://fr.wikipedia.org/wiki/Bulletin_of_the_Atomic_Scientists

Cabanac, G., Labbé, C., & Magazinov, A. (2022). “Bosom peril” is not “breast cancer”: How weird computer-generated phrases help researchers find scientific publishing fraud. *Bulletin of the Atomic Scientists*. <https://thebulletin.org/2022/01/bosom-peril-is-not-breast-cancer-how-weird-computer-generated-phrases-help-researchers-find-scientific-publishing-fraud/>

Eléments biographiques

Guillaume Cabanac est maître de conférences habilité à diriger des recherches (HDR) en informatique à l'Université Toulouse III – Paul Sabatier. Il est membre de l'Institut de Recherche en Informatique de Toulouse (IRIT UMR 5505 CNRS) et siège au Comité national du CNRS en qualité de membre nommé du Conseil scientifique de l'Institut des sciences humaines et sociales (InSHS). Ses travaux interdisciplinaires contribuent à l'analyse de la littérature scientifique, notamment au sein de l'ERC Synergy ‘Nanobubbles’ questionnant le processus d'auto-correction en science. Il développe la plateforme ‘Problematic Paper Screener’⁶ qui signale des milliers d'articles non fiables, pourtant publiés et souvent vendus par les maisons d'édition de premier plan. Cette recherche a été distinguée dans le “Nature's 10”⁷ présentant « dix personnes qui ont aidé à façonner la science en 2021 » selon la revue Nature.

⁶ <https://www.irit.fr/~Guillaume.Cabanac/problematic-paper-screener>

⁷ <https://www.nature.com/immersive/d41586-021-03621-0>