

# Guillaume Cabanac, le Sisyphe de la dépollution de la science

**Chasseurs de fraude — 5/7 —** Cet informaticien de 41 ans a fait de la décontamination de la recherche dévoyée sa spécialité. Et un objet d'étude à part entière

**A**t-on déjà vu un Sisyphe heureux? Oui, et il s'appelle Guillaume Cabanac. Ce professeur d'informatique à l'université Toulouse-III Paul-Sabatier se livre chaque jour, patiemment, à une tâche rébarbative, quasi sans fin, mais qui lui procure plaisir et reconnaissance académique.

Quotidiennement, il plonge dans les milliers d'articles scientifiques publiés. Il ne cherche pas les meilleurs d'entre eux, mais au contraire les pires. Ceux qui sont « pollués » par diverses mauvaises pratiques qui minent la confiance dans la science. Il remonte ainsi dans ses filets des plagiat, des articles écrits avec l'aide de générateurs de texte (bien avant la folie ChatGPT), des articles usurpant des signatures, des productions qui référencent des articles déjà connus comme pollués...

Sa spécialité est d'avoir développé le concept de « phrases torturées », c'est-à-dire des expressions qui ont l'air scientifiques mais que les spécialistes n'utilisent pas. La méthode « Bayes naïve » devient « Bayes crédule », l'intelligence artificielle se transforme en « conscience contrefaite ». Un « algorithme génétique » mute en « calcul héréditaire »... Plusieurs utilisations de ces expressions suspectes, au moins cinq, sont l'indice que l'article a peut-être été plagié ou généré par un algorithme et que des synonymes ont été inventés pour leurrer les logiciels de détection.

Le radar que Guillaume Cabanac a élaboré, le Problematic Paper Screener, recensait, début juillet, plus de 14300 articles épinglés, depuis son lancement en 2021, grâce à neuf « détecteurs » : phrases torturées, références problématiques, recours à des logiciels de génération de textes...

« Je reste scandalisé par ces situations », tranche l'informaticien de 41 ans, qui s'inquiète quand même des conséquences que son travail et sa diffusion pourraient avoir sur la perception que le public a de la science.

« Il faut dire que la plupart des plages sont propres, mais qu'il faut en nettoyer certaines », illustre-t-il. Il s'inquiète aussi de l'étiquette qu'on pourrait lui affubler. « Je ne suis ni un chevalier blanc ni un policier de la pensée. Détective ou limier en tromperie scientifique ou

dépollueur des publications me conviennent mieux. On est là pour corriger la science. »

Et cela le rend plutôt heureux. En juin 2021, quatre collègues l'embarquent dans leur projet récemment financé par le Conseil européen de la recherche (ERC), Nanobubbles. Son but est d'étudier comment la science se corrige et comment, parfois, elle échoue à le faire, lorsque les éditeurs ou auteurs n'agissent pas pour reconnaître les erreurs ou les fautes. Sur Twitter (devenu X), il n'hésite pas à se moquer d'eux et à demander avec humour le versement des frais de publication à des associations. Depuis octobre 2022, il est aussi titulaire d'une chaire à l'Institut universitaire de France sur le thème de la « décontamination » de la littérature scientifique. « Cela légitime cette activité de recherche et me donne des moyens pour la poursuivre », apprécie le chercheur.

## « Pollinisation croisée »

Il est d'autant plus gâté en ce moment que ce travail ingrat lui procure d'autres joies. « Les détectives se structurent! Nous sommes plusieurs à collaborer sur certains cas. On discute, on s'entraide, on échange... J'ai le sentiment d'appartenir à une communauté. J'adore ça. » Il a contribué à cette organisation en lançant, dès 2017, des groupes sur la messagerie Slack qui rassemblent une centaine de personnes environ. Il y a même des représentants d'éditeurs dans certains salons de discussion. Il a également cosigné une lettre ouverte avec une quinzaine d'entre eux pour critiquer la gestion d'une récente affaire d'inconduite par le CNRS.

Il est aussi content de voir la motivation de ses étudiants. « C'est de la pollinisation croisée, souligne-t-il. Pour mon cours sur les bases de données, je me sers de celle qui référence les articles problématiques. Ça motive les étudiants

**« LES DÉTECTIVES  
SE STRUCTURENT! ON DISCUTE,  
ON S'ENTRAIDE, ON ÉCHANGE... »**



OLIVIER BALEZ

qui, en retour, m'aident à repérer de nouveaux cas suspects. » Quand il trouve une phrase tordue, « [il] contacte les auteurs et leur demande s'ils peuvent [lui] citer des références, des livres dans lesquels ces phrases se trouvent » : « Je m'inquiète qu'ils fassent référence, peut-être sans le savoir, à des articles déjà repérés comme problématiques. »

## Les bas-fonds de la recherche

Les moissons sont, hélas, bonnes. Les plus grosses ont même touché des éditeurs reconnus, comme Elsevier ou Springer, pour chacun environ 400 articles, repérés en 2021. Pour l'occasion, Guillaume Cabanac s'était associé à deux auteurs. L'un, Cyril Labbé, professeur à l'université Grenoble-Alpes en informatique, l'avait éveillé à ces questions grâce à un canular. En 2010, celui-ci avait montré comment un chercheur imaginaire, Ike Antcare (prononcer « I can't care » pour « je ne peux pas m'en soucier »), était devenu la vedette de son domaine avec de faux textes. L'autre, Alexander Magazinov, est un amateur perspicace, qui nourrit le blog d'un autre chasseur de fraudes, Leonid Schneider.

Ces succès dans les bas-fonds de la recherche ont, pour Guillaume Cabanac, des racines

plus anciennes. En 2015, il publie dans une revue de Wiley un article sur Library Genesis, une base de données d'articles payants, piratés et mis en ligne gratuitement. Ce succès lui a apporté une première reconnaissance et « cela [l]a conforté dans [s]on idée que les articles peuvent être des objets de recherche à part entière ».

Si le détective est satisfait de ce travail de nettoyage, il veut aussi faire du préventif. Il ne prétend pas que les malhonnêtetés disparaîtront, mais il voudrait qu'au moins des articles paresseux, fabriqués, plagiés ne sortent plus. Il faut donc de la formation, de la collaboration, des éditeurs proactifs...

Problème, de nouveaux outils de génération de textes comme les grands modèles de langage en intelligence artificielle montrent tout leur potentiel néfaste. Guillaume Cabanac réfléchit à des moyens pour, malgré ces logiciels, repérer de nouveaux abus en testant les différences entre les articles suspects et le corpus du domaine auquel ils appartiennent. Sisyphe reste heureux. ■

DAVID LAROUSSERIE

La semaine prochaine Clara Locher, à l'affût des liaisons dangereuses

# Comment l'onagre écoute les insectes voler

**Plantes de génie — 5/7 —** Cette fleur jaune peut détecter ses pollinisateurs à distance et communiquer avec eux pour les attirer

**P**ieds nus dans les dunes, l'estivalier de Marseillan Plage, tout proche de Sète (Hérault), peut ne pas la remarquer. Touffue ou poussant en grappes qui rasent le sable, l'onagre de Drummond a pourtant une qualité extraordinaire que ne laisse pas deviner son apparence un brin quelconque. Cette fleur jaune peut détecter à distance ses principaux pollinisateurs. Plus précisément, ses pétales vibrent au son des ailes des abeilles. En réaction, la fleur produit un nectar adouci, 20 % plus sucré. Un mécanisme d'attraction, qui accroît ses chances de reproduction, activé en trois minutes chrono!

En France, la station balnéaire qui jouxte l'étang de Thau est le seul endroit répertorié – par l'Inventaire national du patrimoine naturel (INPN) – où l'*Oenothera drummondii* pousse de façon sauvage. « La fleur a dû s'échapper d'un jardin », explique Guillaume Fried, botaniste et chercheur à l'Agence nationale de sécurité sanitaire de

l'alimentation, de l'environnement et du travail (Anses), qui analyse sur le territoire les risques de développement des plantes exogènes. « En France, elle est naturalisée, c'est-à-dire qu'elle se maintient depuis quelques années sans proliférer. On peut donc encore l'acheter pour un usage privatif. Par contre, dans le sud de l'Espagne, elle est déjà considérée comme invasive », précise le chercheur.

Originaire du Texas, où le botaniste écossais Thomas Drummond l'a découverte en 1833, cette onagre est « un envahisseur efficace qui a colonisé de nombreux systèmes dunaires dans le monde entier au cours du dernier siècle », précisent des universitaires espagnols et mexicains dans une étude publiée en novembre 2019. La mappemonde proposée par l'INPN permet de visualiser les plausibles voyages de ses graines. Traversant mers et océans, la plante s'est introduite de façon pointilliste au Pays de Galles, en Suède, au Pérou, en Chine, en Afri-



A Marseillan (Hérault), en octobre 2014. GUILLAUME FRIED

que du Sud, en Egypte, en Australie et encore en Israël.

Dans ce dernier pays, l'onagre, qui fleurit à foison sur les plages, inspire particulièrement les chercheurs. En 1986, des travaux, menés par le professeur en botanique

Dan Eisikowitch, de l'université de Tel-Aviv, ont permis de comprendre que la fleur avait la capacité de faire savoir, par la forme et la couleur de ses pétales, si elle avait déjà été vidée de son pollen. L'équipe a ainsi mis au jour la manière dont

l'onagre aiguillait les pollinisateurs « pour qu'ils économisent de l'énergie, trouvent les bonnes fleurs et ne quittent pas la zone sans l'avoir totalement exploitée », écrit-elle dans le *Botanical Journal of the Linnean Society*.

## Même longueur d'onde

Depuis la même université, une trentaine d'années plus tard, la professeure Lilach Hadany, théoricienne de l'évolution, a démontré en 2019 que l'onagre pouvait entendre ses pollinisateurs et leur répondre pour mieux les attirer. Pour le prouver, elle et son équipe ont soumis la fleur placée en laboratoire à cinq sons différents : le silence – sous une cloche de verre –, des enregistrements d'une abeille à une distance de 10 centimètres en vol stationnaire (fréquence : de 0,2 à 0,5 kilohertz), et des sons générés par ordinateur dans des fréquences basses (de 0,05 à 1 kilohertz), moyennes (de 34 à 35 kilohertz) et hautes (de 158 à 160 kilohertz). Dans le silence et sous

hautes fréquences, la plante n'a pas réagi, mais, sous de basses fréquences, que le son soit produit artificiellement ou par les abeilles, le nectar de la plante s'est sucré.

Lilach Hadany avait l'intuition que les « oreilles » de la plante, dont la corolle a des airs de mini-antenne parabolique, étaient ses pétales. A l'aide d'un vibromètre laser – qui mesure d'infimes mouvements –, l'équipe a pu confirmer que pétales et abeilles, à proximité, entraient en résonance et vibraient sur la même longueur d'onde. Par contre, si certains pétales manquent, l'oreille végétale fonctionne moins bien.

De nombreuses questions restent en suspens mais l'onagre de Drummond peut déjà s'enorgueillir d'avoir ouvert la voie à un nouveau champ de recherche : la phytoacoustique. ■

LAURE BELOT

La semaine prochaine L'androsace du Dauphiné, alpiniste menacée