



HAL
open science

Fake science : un (vrai) panorama

Nicolas Jardin, Norbert Philippe, Arnaud Dubos, Marion Callet

► **To cite this version:**

Nicolas Jardin, Norbert Philippe, Arnaud Dubos, Marion Callet. Fake science : un (vrai) panorama. Doctorat. Ecully, France. 2022, pp.7. hal-03828647

HAL Id: hal-03828647

<https://hal.science/hal-03828647v1>

Submitted on 17 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fake science et fraudes scientifiques

La crise sanitaire, par son ampleur, a mis le doigt sur la problématique fondamentale de la confiance des citoyens et de la société dans la science.

La question, qui n'est pas nouvelle, nécessite de réaffirmer les principes nécessaires à la bonne marche de la science : intégrité, éthique, reproductibilité des résultats, ouverture, transparence.

Elle nécessite aussi de savoir ce qui est à combattre : **de quoi parle-t-on quand on parle de « fake science » ?**



Au sens large, la fake science, qu'on peut traduire en français par pseudo-science ou malscience, désigne toutes les affirmations qui se présentent comme scientifiques mais qui ne respectent pas les règles et bonnes pratiques de la méthode scientifique.

La fake science peut être produite et relayée par n'importe quel citoyen : des professionnels de la fraude, des influenceurs, des journalistes, des politiques... mais aussi des scientifiques.

De l'approximation à la fraude : 50 nuances de méconduite scientifique

Fraude	Pratiques discutables en recherche	Méconnaissance méthodologique
<ul style="list-style-type: none">FabricationFalsificationPlagiat	<ul style="list-style-type: none">« Torture » ou « massage » des donnéesChangement du critère de jugementChoix sélectif / omission de donnéesRéférences erronéesChangement de tests statistiquesP-Hacking (Hypothesizing After Results are Known)Manipulation d'imagesPaternité des articles : conflits d'auteursÉtudes animales trompeusesNon-publication de recherches financéesRésumés, communiqués de presse embellis...	<ul style="list-style-type: none">Méthodes « faibles »Méthodes inappropriéesÉchantillon trop faibleErreurs statistiquesPas de recherche documentaire avant la rechercheNon respect de réglementations

Intentionnel (rouge) / Non intentionnel (vert)

Figure adaptée par Hervé Maisonneuve de Seror et Ravaud, « Embellissement des données : fraude a minima, incompetence ou mélange des deux », La presse médicale, 2012.

Quand on parle de **méconduite scientifique**, le spectre est large, allant de l'approximation ou de la méconnaissance méthodologique, jusqu'à la fraude volontaire.

Entre les deux se décline tout un nuancier de pratiques que l'on peut regrouper sous le terme de « **pratiques de recherche douteuses** » (*questionable research practices*)¹.

Les exemples de méconduite les plus médiatiques sont en général des cas de fraude scientifique, définie comme « une violation sérieuse et intentionnelle dans la conduite d'une recherche et dans la diffusion de résultats² » et traditionnellement décomposée en trois ensembles :

- **Falsification des données** : les données obtenues sont manipulées de façon à obtenir le résultat attendu. Charles Babbage parlait dès 1830 de données « taillées » ou « cuisinées » : les données gênantes n'allant pas dans le sens de la démonstration sont éliminées afin d'aboutir à un résultat positif.

- **Fabrication** : les données sont créées de toutes pièces.

- **Plagiat** : désigne l'appropriation des idées, des résultats ou des mots d'une autre personne, sans lui en donner correctement crédit. Il peut prendre diverses formes :

- **reproduction** de texte ou d'illustrations provenant d'une autre étude sans en mentionner la source ;
- **vol d'idées** : exemple d'un relecteur reprenant à son compte des idées exposées dans un article dont il a refusé la publication ;
- « **oubli** » de mentionner un des auteurs de l'article...

1 : Concept popularisé par John, Loewenstein et Prelec dans « Measuring the prevalence of questionable research practices with incentives for truth telling », Psychological science, 2012.

2 : Letellier, « Sur l'intégrité de la recherche : quelques considérations éthiques sur l'organisation et les pratiques de recherche », Revue Prétextaine, 2011



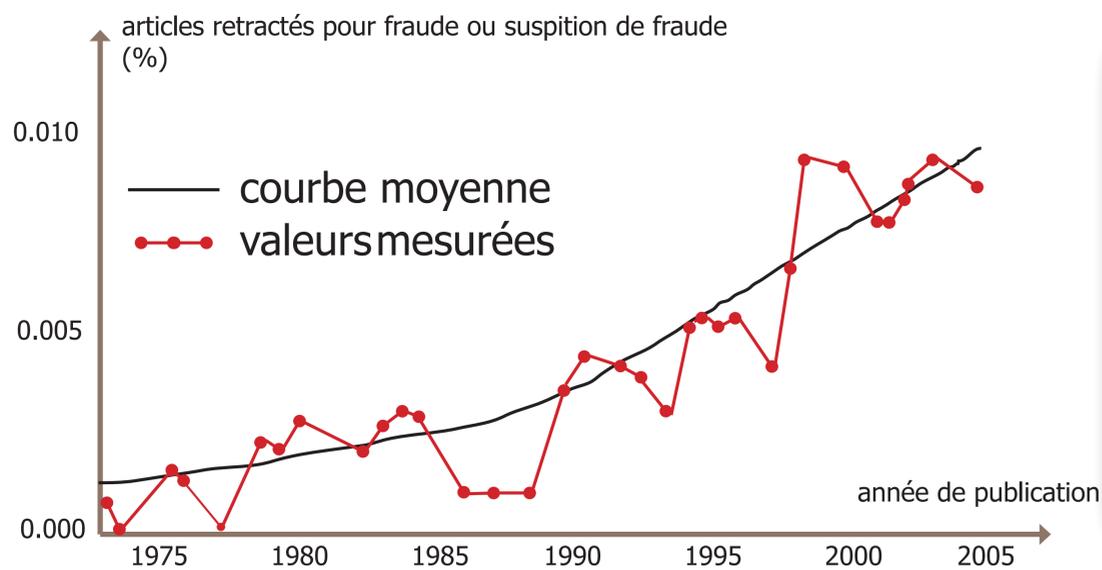
Fake science et fraudes scientifiques

Fraudes scientifiques : quelle ampleur et quelles conséquences ?

Le développement des bases de données bibliographiques (PubMed, Web of Science, Scopus...) a permis d'évaluer l'ampleur du phénomène, en recensant les articles scientifiques publiés et rétractés, c'est à dire invalidés.

La fraude reste marginale rapportée au nombre de publications indexées : 1 à 2 % des quelques 1 500 000 articles publiés chaque année dans le monde seraient frauduleux.¹ Mais elle tendrait à gagner du terrain : une étude a montré qu'en biomédecine, une des disciplines les plus touchées, le taux de rétractation d'articles pour fraude avait été multiplié par 10 entre 1975 et 2010.

Enquête auprès des chercheurs



Evolution des rétractations d'articles pour fraude - Source Fang et al. - 2012

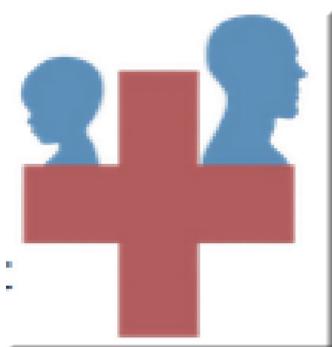
Une fraude réussie étant une fraude non détectée, les enquêtes réalisées auprès des chercheurs permettent d'avoir une idée plus précise de l'ampleur du phénomène. En 2009, une enquête menée auprès de chercheurs anonymes indique que 2 % des chercheurs admettent avoir déjà falsifié ou fabriqué des résultats, mais 14 % affirment avoir relevé ces pratiques chez des collègues² !

Ces fraudes et méconduites scientifiques, parfois couvertes médiatiquement, ont plusieurs conséquences néfastes :

- **La perte de confiance des citoyens et de la société dans la science** : la crise sanitaire l'a douloureusement rappelé. Un article écrit par un chercheur, travaillant dans un laboratoire rattaché à une université ou une grande école apporte une caution scientifique à des thèses qui peuvent ensuite être très rapidement relayées dans les médias ou sur les réseaux sociaux.

- **Un effet boule de neige sur la connaissance scientifique** : la science étant cumulative, des résultats frauduleux sont susceptibles de polluer d'autres articles les reprenant, et de se propager ainsi au sein de la communauté scientifique, jusqu'à ce que la fraude soit découverte.

- **Un gaspillage des fonds publics** : les résultats frauduleux, non reproductibles ne permettent pas d'espérer de retombées positives pour la société qui a financé ces recherches.



- **Les conséquences peuvent être graves** si des études frauduleuses influent sur des décisions publiques, notamment dans le domaine de la santé.

Un des exemples les plus connus est celui de la recrudescence des cas de rougeole au Royaume-Uni suite à la publication en 1998 par Andrew Wakefield d'une étude faisant un lien entre vaccin ROR (rougeole-oreillons-rubéole) et autisme³

Il a depuis été démontré que l'article reposait sur des dossiers médicaux falsifiés.

1 : Barde, Peiffer-Smadja, de la Blanchardière, « Fraude scientifique : une menace majeure pour la recherche médicale », La revue de médecine interne, 2020.

2 : Fanelli, "How many scientists fabricate and falsify research ? A systematic review and meta-analysis of survey data", PLoS One, 2009.

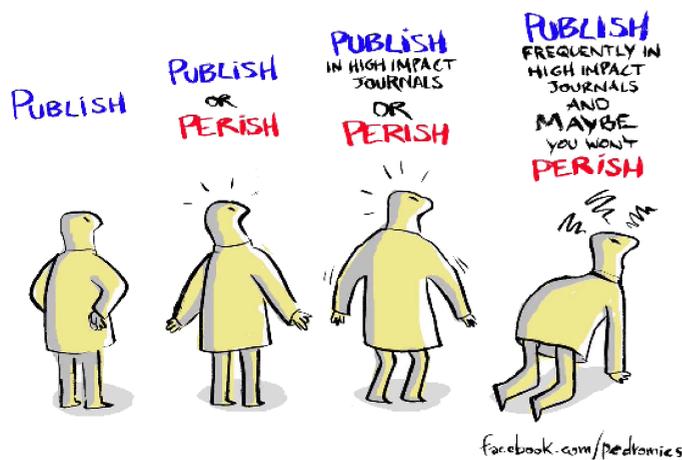
3 : Article rétracté par The Lancet en 2010 : Wakefield, "Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children", Lancet, 1998.

Garde-fous et remèdes contre la fake science

Pression financière sur les laboratoires de recherche, compétition entre universités et entre chercheurs, budgets alloués et progression de carrière liés au nombre de publications, valorisation exclusive des résultats positifs... Ces éléments de contexte poussent les chercheurs à publier de plus en plus et de plus en plus vite selon l'adage *Publish or perish*, favorisant les tentations de prendre parfois des libertés avec l'intégrité scientifique. Face à cela, des garde-fous ont été mis en place, et des réflexions sont menées pour lutter contre ces pressions diverses qui peuvent pousser aux méconduites scientifiques.

Une réaffirmation de l'intégrité scientifique

THE EVOLUTION OF ACADEMIA



Dessin Pedromics [Pedro Velica],
"The evolution of academia", Facebook.

« L'intégrité scientifique est l'ensemble des valeurs et des règles qui garantissent l'honnêteté et la rigueur de la recherche et de l'enseignement supérieur. Elle est indispensable à la cohésion des collectifs de recherche et à l'entretien de la confiance que la société accorde à la science » (HCERES).

L'intégrité scientifique concerne l'ensemble des étapes du processus de recherche, de la recherche de financement à la publication, en passant par le travail de recherche, la communication, la propriété intellectuelle, l'évaluation et l'expertise...

Quelques repères :

2015

8 opérateurs de recherche, dont le CNRS, signent la 1ère Charte nationale de déontologie des métiers de la recherche.

2017

- Lancement de l'Office Français de l'Intégrité Scientifique (OFIS) ;
- Publication du guide « Pratiquer une recherche intègre et responsable » (CNRS-CPU).

2016

La formation à l'éthique de la recherche et à l'intégrité scientifique des doctorants devient obligatoire.

2018

- Code de conduite européen pour l'intégrité en recherche ;
- Lancement du Plan national pour la science ouverte.



Repenser les critères d'évaluation de la recherche

De nombreuses initiatives, issues du monde de la recherche, visent à repenser ses critères d'évaluation.

C'est ainsi le cas de la **Déclaration de San Francisco sur la recherche** (DORA, 2013), qui recommande entre autres :

- de ne pas utiliser d'indicateurs basés sur les revues (comme le facteur d'impact), mais bien sur les articles eux-mêmes ;
- de tenir compte de la valeur et de l'impact de tous les résultats de recherche, y compris les jeux de données et les logiciels, en plus des publications scientifiques ;
- d'envisager un large éventail de mesures d'impact, y compris des indicateurs qualitatifs sur les retombées des travaux, comme leur influence sur les politiques et les pratiques.



Garde-fous et remèdes contre la fake science

Pour une science plus ouverte

Le mouvement de la science ouverte, en visant « **à construire un écosystème dans lequel la science sera plus cumulative, plus fortement étayée par des données, plus transparente, plus rapide et d'accès universel** », apporte des réponses à la perte de confiance d'une partie de la société dans la science.

Dans le cadre des sciences participatives, les citoyens sont même associés à la collecte ou l'analyse des données.

Une science plus ouverte permet de prévenir et de lutter contre la fraude en favorisant la transparence des processus, méthodes, données, codes... qui ont permis d'aboutir à des résultats. Elle permet également à d'autres chercheurs d'essayer de reproduire ces résultats pour vérifier leur solidité.



Composantes de la science ouverte

« Plan national pour la science ouverte », Ouvrirlescience.fr, 2018.

L'exemple du **Lancet Gate** l'a démontré : un article publié dans le Lancet dénonçait l'inefficacité de l'hydroxychloroquine contre la COVID-19.

Mais de nombreuses incohérences ont alerté. Quand les scientifiques ont demandé accès aux données brutes soutenant l'article, ils se sont heurtés à une fin de non recevoir.

Il est finalement apparu que la start-up qui avait récolté les données - et dont un des auteurs de l'article était PDG - était une société fantôme, dont la rédactrice scientifique était une spécialiste... de science-fiction.

Temps médiatique VS temps scientifique

Le temps politique et le temps médiatique ne sont pas ceux de la science : la **méthode scientifique** vise à produire des connaissances fiables et vérifiées, ce qui demande du temps.

Or, les décideurs peuvent demander des réponses urgentes, et les informations circulent de plus en plus rapidement, rendant difficile leur vérification en amont, même dans les médias traditionnels.

Le risque de voir le travail et les résultats des chercheurs déformés, tronqués, surinterprétés ou instrumentalisés est alors réel.

La crise sanitaire a donné de nombreux exemples de traitements de l'information qui dénotaient une méconnaissance des réalités scientifiques.

À titre d'exemple, la **controverse scientifique** est un élément essentiel pour la progression des connaissances, mais son exercice dans les médias peut être mal compris et perçu comme un aveu d'ignorance.



« How science reporting works », smbc-comics.com, 2009.

1 : Unesco. Composantes de la science ouverte. 2020.

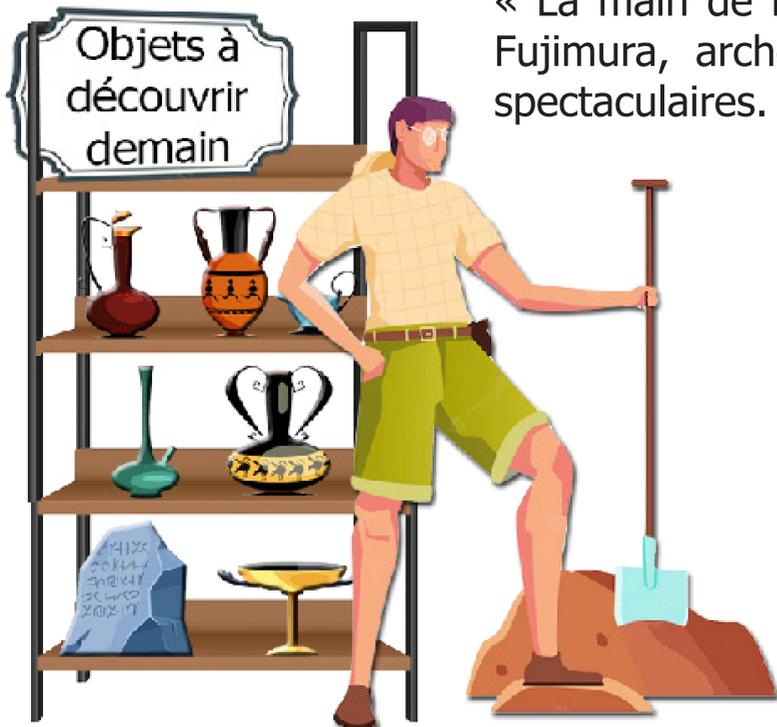


Contenu : N. JARDIN et N. PHILIPPE
Mise en forme : M. CALLET et A. DUBOS

Le worst of de la fraude scientifique

Ces dernières années, des cas spectaculaires de fraudes scientifiques ont été relayés par les médias. S'il convient de souligner que les méconduites scientifiques les plus courantes sont plus discrètes, nous vous proposons ci-dessous un petit bêtisier des fraudes scientifiques à travers quatre exemples.

L'archéologue qui enterrait avant de déterrer



« La main de Dieu », « l'excavateur divin »... ainsi était surnommé Schinichi Fujimura, archéologue japonais dont les découvertes étaient toujours plus spectaculaires.

Auteur de quelque 180 fouilles fructueuses à travers le Japon, il annonce en octobre 2000 avoir découvert des vestiges vieux de 570 000 ans.

Problème : quelques jours plus tard, un quotidien japonais publie des photographies de l'archéologue en train d'enterrer lui-même les trésors que lui et son équipe s'apprêtaient à découvrir... Fujimura fut démis de ses fonctions après être passé aux aveux, reconnaissant avoir enterré 61 des 65 pièces découvertes, issues de sa propre collection.

Ike Antkare, ce grand chercheur... qui n'existait pas

En octobre 2010, un mystérieux chercheur nommé Ike Antkare apparaît sur Google Scholar, avec de nombreux articles et un h-index* de 90, bien au-dessus de ses plus éminents collègues... Aurait-on affaire à un nouvel Einstein, apparu soudainement sur le devant de la scène scientifique ?

En réalité, derrière ce pseudonyme évocateur (Ike Antkare = « I can't care » = « je m'en fiche ») se cachait Cyril Labbé, chercheur en informatique à l'Université de Grenoble. En créant cet avatar, puis en publiant sur sa page personnelle du site de son université des articles scientifiques générés aléatoirement et se citant les uns les autres, il est devenu en quelques semaines l'un des plus grands scientifiques de son temps d'après les indicateurs de Google Scholar. Il a surtout atteint son réel objectif : mettre en lumière la toute relative fiabilité de ces indicateurs de performance, surtout quand un large moissonnage comme celui effectué par Google Scholar ne permet pas de garantir la fiabilité des articles indexés.



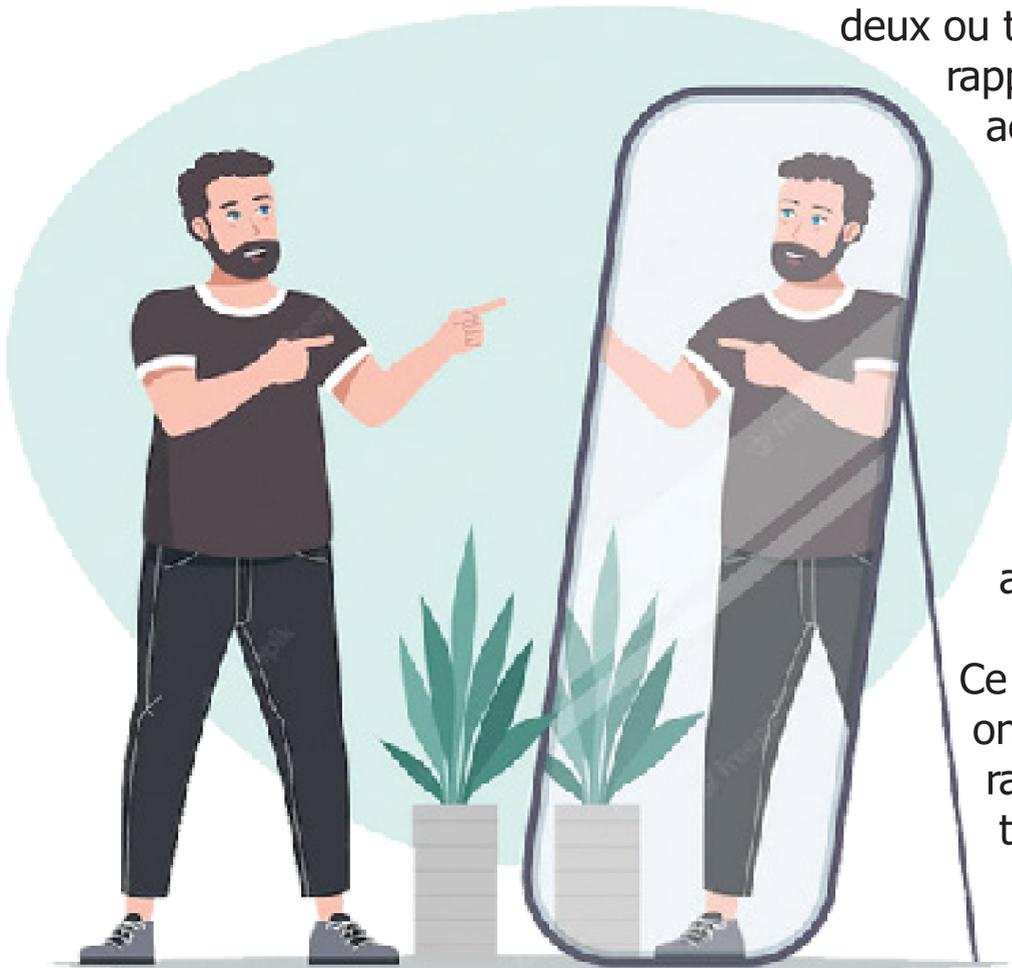
* Score traduisant le nombre de publications et de citations obtenues par un chercheur.



Le worst of de la fraude scientifique

Le chercheur qui évaluait ses propres publications (et les trouvait pas mal)

La relecture par les pairs *peer-reviewing* est le système qui permet de garantir, autant que possible, la validité scientifique des articles publiés. Chaque article soumis à une revue scientifique est évalué par deux ou trois experts du domaine concerné, qui rédigent un rapport d'évaluation indiquant si l'article peut être accepté pour publication et si oui, avec quelles modifications.



Il peut arriver que le comité éditorial d'une revue demande à l'auteur de suggérer des noms de relecteurs. Or quand Hyung-In Moon, chercheur spécialisé dans l'étude des plantes, suggérait des relecteurs dont les noms étaient usurpés ou inventés, il donnait de faux e-mails de contact qu'il avait lui-même créés...

Ce n'est pas tant les évaluations, toujours positives, qui ont attiré l'attention sur le stratagème, mais l'incroyable rapidité avec laquelle les rapports d'évaluation tombaient : en général moins de 24h.

Le recordman du monde des rétractations d'articles

Une rétractation d'article scientifique est l'action par laquelle un article publié est invalidé. Il reste visible en ligne, mais est accompagné d'une notice rédigée par l'éditeur, expliquant les raisons de la rétractation. Un article peut être rétracté pour fraude, mais aussi pour des erreurs de bonne foi.

Mais lorsqu'un chercheur compte 183 articles rétractés, et que, dans 171 cas au moins, il a été établi que les données avaient été fabriquées, le doute s'amenuise...

Le site **Retraction Watch**, qui effectue une veille sur les articles rétractés, a ainsi raconté l'histoire de Yoshitaka Fujii, anesthésiste japonais, dont les données étaient, selon le trio de chercheurs qui a lancé l'alerte, « *incredibly nice* » (ce qui, en science, n'est pas un compliment).



Dans un contexte concurrentiel entre chercheurs et face à des évaluations largement basées sur des critères quantitatifs, des professionnels de la fraude ont bien compris qu'ils tenaient là un substrat favorable sur lequel prospérer.

C'est ainsi que de faux éditeurs et de fausses revues scientifiques se sont développés, faisant miroiter aux chercheurs la garantie de publications rapides et en accès ouvert - accessibles gratuitement à tous - contre quelques centaines d'euros.

Ces revues prédatrices utilisent des stratagèmes de plus en plus sophistiqués pour tromper les chercheurs et s'assurer de revenus confortables, en publiant des résultats de recherche hors de tout cadre garantissant leur fiabilité.

Comment ça marche ?

Editeur prédateur



Chercheur

Envoi de spams flatteurs aux chercheurs



Une publication sur le modèle auteur-payeur



Révision par les pairs expéditive ou inexistante



Publication sur le site frauduleux de la revue

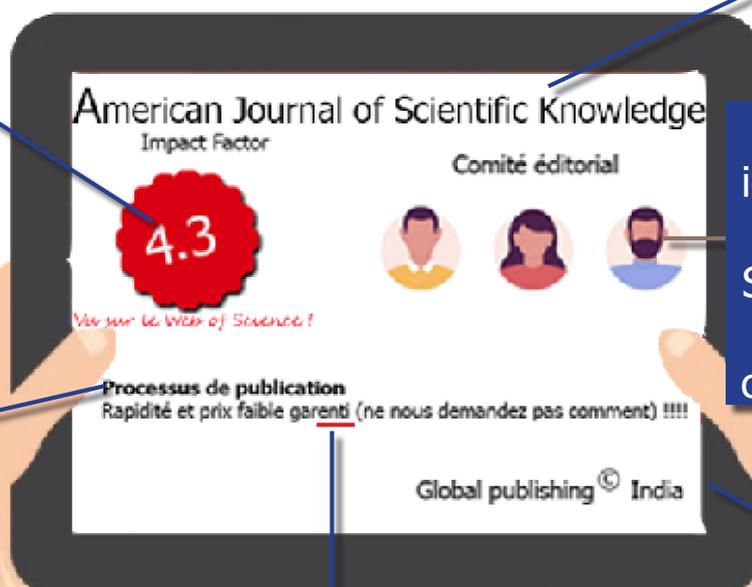
Quelles conséquences pour le chercheur piégé ?

- Impact négatif sur sa réputation et sur celle de son institution.
- Article non pris en compte dans les évaluations.
- Perte d'argent pour lui ou son institution.
- Perte de ses droits d'auteur ; impossibilité de republier l'article ailleurs.
- Risque de devoir rétracter son article.

Comment reconnaître une revue prédatrice ?

- Faux facteur d'impact ;
- Indexation bidon dans les grandes bases de données

Titre reprenant celui d'une vraie revue en le modifiant légèrement



Comité éditorial difficile à identifier ou que l'on retrouve dans plusieurs revues. S'il s'agit de vrais chercheurs, ils ne sont même pas au courant de leur participation !

- Délais de *peer-reviewing* et de publication très (trop) courts ;
- Prix dérisoire des frais de publication (APC) ;
- Manque de transparence sur le processus de publication

Fautes d'orthographe et de grammaire

L'origine géographique de la revue ne correspond pas au titre

Fake science : un exemple français



Olivier Voinnet

Médaille d'argent CNRS 2009

Chercheur internationalement reconnu en biologie médicale suite à ses travaux sur l'ARN



PubPeer est une plate-forme destinée à soumettre aux feedbacks de la communauté scientifique des articles publiés. Les commentaires postés, pour la plupart anonymes, peuvent porter sur des suggestions d'amélioration des articles, mais aussi sur des soupçons de méconduite scientifique.



Dès 2015 apparaissent sur PubPeer des soupçons quant à la crédibilité des résultats d'un article publié par O. Voinnet en 2003, aboutissant à sa rétractation.

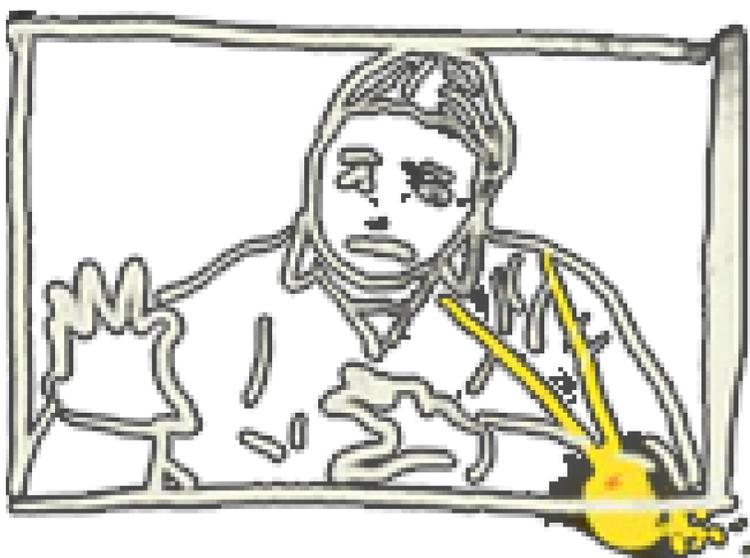
Ecole Polytechnique Fédérale de Zurich

Le CNRS puis l'EPFZ, employeurs du chercheur, enquêtent. 20 publications sur 32 expertisées sont entachées de manipulations délibérées de figures : modifications et embellissement d'images, duplications d'éléments, légendes erronées...



Exclusion du CNRS pour 2 ans. 8 articles sont rétractés de la presse scientifique.

POURQUOI ? Publish or perish



« c'était dû à la RAPIDITE et non pas à l'erreur. Je n'avais pas le temps, je montais mes figures n'importe comment et je m'en fichais complètement. »

La commission d'enquête de l'EPHZA indiqué qu'Olivier Voinnet a lui-même « admis que certains articles avaient été publiés trop vite, dans un environnement hautement compétitif ».