



Decontamination of the scientific literature

Guillaume Cabanac

► To cite this version:

Guillaume Cabanac. Decontamination of the scientific literature. Institut Universitaire de France. 2022. hal-03835147

HAL Id: hal-03835147

<https://ut3-toulouseinp.hal.science/hal-03835147>

Submitted on 31 Oct 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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Decontamination of the scientific literature

Guillaume Cabanac

Submitted: November 2, 2021 · Granted: May 2, 2022 · Preprinted: October 28, 2022

Abstract Research misconduct and frauds pollute the scientific literature. Honest errors and malevolent data fabrication, image manipulation, journal hijacking, and plagiarism passed peer review unnoticed. Problematic papers deceive readers, authors citing them, and AI-powered literature-based discovery. Flagship publishers accepted hundreds flawed papers despite claiming to enforce peer review. This application ambitions to decontaminate the scientific literature using curative and preventive actions.

Keywords research integrity · scientific literature · whistleblowing · errors · fraud

1 Context: Fraud and misconduct undermine trust in science

In 2020 the scientific community reached a new world record despite the global COVID19 pandemic. Scholarly knowledge production peaked at [5 million peer-reviewed articles](#) as per the Dimensions bibliographic database ([Herzog, Hook, & Konkiel, 2020](#)). Year after year, research production grows in terms of publication counts ([Soete, Schneegans, Eröcal, Angathevar, & Rasiah, 2015](#), p. 36) and geographic footprints ([Maisonobe, Jégou, & Cabanac, 2018](#)). The ‘publish or perish’ atmosphere, which is already decades-old ([Cabanac, 2018](#)), intensifies when scientific papers once viewed as *knowledge* units are increasingly considered as *accounting* units ([Gingras, 2020](#)). Some institutions—countries, even—have implemented financial incentives to reward authors who manage to publish in the most highly-regarded journals ([Chen, 2019](#); [Lin, 2013](#)). Such incentives contravene the ‘disinterestedness’ norm of the Ethos of Science ([Merton, 1942](#)) and they indirectly fostered fraud and misconduct to ‘game the metrics’ ([Biagioli & Lippman, 2020](#)) with documented cases of:

- ‘Paper mills’ as industrialised ghostwriting where companies churn pseudo-scientific papers and sell authorship positions ([Else & Van Noorden, 2021](#); [Oransky, 2021b](#)).

G. Cabanac

University of Toulouse, Computer Science Department, IRIT UMR 5505 CNRS, 31062 Toulouse, France
Institut Universitaire de France (IUF), 75005 Paris, France

E-mail: guillaume.cabanac@univ-tlse3.fr

ORCID: [0000-0003-3060-6241](https://orcid.org/0000-0003-3060-6241)

- Publication with predatory publishers implementing weak peer-review (Faulkes, 2021; Grudniewicz et al., 2019).
- Hijacked journals fooling authors by replicating logos and names of legitimate journals (Abalkina, 2021).
- Identity theft to impress editors/reviewers with prestigious affiliations (Barbash, 2014) and fool editorial assessment with forged researcher identifiers (Teixeira da Silva, 2021).
- Data fabrication as in the Surgisphere scandal during the pandemic (Piller, 2020).
- Image manipulation, especially in the health sciences (Shen, 2020).
- Undeclared conflicts of interests casting doubts on authors' motivations and their results ("Confronting conflict of interest [Editorial]", 2018; Lewison & Sullivan, 2015).
- Computer-generation of nonsensical text to pad papers or even generate entire studies (Cabanac & Labbé, 2021).
- Plagiarism of published articles (Pupovac, 2021) using thesaurus-based paraphrasing (Cabanac, Labbé, & Magazinov, 2021).

It is of critical importance to find and remove such pollution from the literature, as erroneous publications potentially deceive:

- *Readers* whose expertise in the field might not be high enough to detect errors. Readers turned authors may even propagate the errors when citing flawed papers.
- *Meta-researchers* who perform systematic reviews of the literature that may contain flawed papers. What if the evidence used to recommend a medical treatment is based on an inaccurate study? Preprints proved challenging for living systematic reviews (Oikonomidi et al., 2020) and so are erroneous publications.
- *AI-powered software* performing literature-based discovery (Bruza & Weeber, 2008) and writing literature reviews automatically (Beta Writer, 2019).

While retractions of problematic publications are on the rise (Abritis, Marcus, & Oransky, 2020; Brainard & You, 2018; Sharma, 2021) the public is entitled to ask: How many problematic papers go unnoticed? This sadly contributes to the 'reproducibility crisis' in various research areas (Baker, 2016), which affects the trust people place in science.

In October 2021, the Institut Universitaire de France (IUF) organised its annual congress titled *Post-truth? Credibility of scientific research in a time of "alternative facts"* to foster research on this concerning issue. The talks I attended laid a dark landscape of science in our post-truth era; they convinced me to apply to the IUF Chair programme.

2 This IUF application: Decontamination of the scientific literature

During my 5-year tenure as Junior IUF member, I plan to tackle the issue of pollution staining the scientific literature, a current concern shared with the *Office Français de l'Intégrité Scientifique* (HCÉRES-OFIS, 2021). My plan is twofold. Its *curative* part deals with the problematic papers that are already published: we need to find them, report them, and get them down. Its *preventive* part ambitions to anticipate the new forms of misconduct and react as soon as spoiled literature pops up.

2.1 Curative approach for decontaminating the scientific literature

The pollution staining the scientific literature is here to stay until responsible scientists or publishers spot it and take action. Individually or as team member, I have initiated three

concrete actions to wipe the literature from erroneous studies. The IUF Chair would contribute to upscale and enhance these ongoing endeavours that already yielded substantial results.

2.1.1 Detection and correction of the pollution by paper mills in oncology

In 2014, a Professor of molecular oncology at the University of Sydney noticed errors in the genetic sequences reported in articles published by a variety of journals. These papers would claim that a given DNA sequence 1) targets human gene X or 2) targets none of the human genes. These claims form the basis of gene knockdown experiments and any inaccuracy invalidates any results presented. Meticulous readers would check these claims with the [BLAST](#) search tool hosted by NCBI. This Prof. Jennifer Byrne did, only to find hundreds of erroneous claims in *de facto* erroneous papers. Her sleuth work gained high recognition in 2017: The *Nature* journal listed her in the Top 10 of people who mattered that year ([Phillips, 2017](#)).

Prof. Byrne teamed up with Dr. Cyril Labbé, an Associate Professor from Grenoble to build the [Seek&Blastn software](#) that automates the screening of DNA sequences extracted from papers ([Labbé, Grima, Gautier, Favier, & Byrne, 2019](#)). I joined the Byrne–Labbé collaboration and we designed an Information Retrieval task and evaluation benchmark to assess the effectiveness of error detection in papers reporting genetic sequences ([Labbé et al., 2020](#)). We screened thousands of papers and reported the problematic ones to editors-in-chief and publishers. Problems relate to unsupported claims (targetting/non targetting DNA sequences) and untraceable source of genetic materials used (e.g., unknown Hollybio company named as supplier of biologic material). Most suspect papers follow a common template (e.g., study of the effect of gene X on organ/condition Y) that is typical of the ‘paper mills’ output ([Else & Van Noorden, 2021](#)). Despite our team’s continuing efforts, formal retractions are long to come—when they do come ([Byrne et al., 2021](#)).

Our latest research revealed the extent of problematic papers in two oncology journals that we screened cover-to-cover thus checking 13,700 nucleotide sequences in 3,400 papers. The 21% error-rate is unacceptably high for this scholarly literature. In addition, the positive reception of these erroneous papers (17k citations to 712 papers) suggests citation manipulation at scale ([Park et al., 2021](#)). This alarming result has been profiled in the *News* section of *Nature* ([Else, 2021a](#)) and we currently analyse the suspect citation networks to hopefully uncover any citation cartels.

2.1.2 Detection and reporting of hijacked and hacked scientific journals

Predatory publishers are known to sell quick time-to-market with little to no peer review ([Grudniewicz et al., 2019](#)). The reputable titles and logos of established journals get hijacked when scammers copy them to create fake lookalikes ([Abalkina, 2021](#)). These fool the inattentive—naïve—authors who believe they submitted to a reputable venue. Crooked authors also submit weak or even nonsensical computer-generated papers knowingly.

Flagship publishers were considered immune to nonsensical submissions. It was especially true for ‘elite’ journals with Impact Factors—20% only have one. And yet, we found that an Elsevier journal with Impact Factor, *Microprocessors and Microsystems*, had been compromised in 2021. It had published no less than 400 problematic articles when we released our whistleblowing study reporting tortured phrases and shrinking durations of editorial assessments ([Cabanac et al., 2021](#)). Elsevier acknowledged the errors and asked experts to reassess all these suspect articles ([Else, 2021c](#); [Marcus, 2021a, 2021b](#)).

We also contributed to uncover a similar integrity breach at Springer’s *Arabian Journal of Geosciences*. This journal with Impact Factor had published more than 400 absurd papers when we reported this case to the public (Oransky, 2021a, 2021c). Analysing the timelines of editorial assessments I highlighted an unexpectedly steep time-to-market shrinkage starting in 2021.¹ Responsible publishers should monitor and investigate such sudden changes in their production to prevent other nonsensical ‘bubbles’ bursting after hundreds problematic papers were published.

2.1.3 Detection and reporting of nonsensical algorithmically generated articles

Highly-regarded publishers in engineering, such as ACM and IEEE, have published and sold meaningless computer-generated papers. We designed an algorithm to comb the literature for fraudulent papers generated with probabilistic context-free grammars such as *SCIgen* and *Mathgen*. We found 262 such papers, 197 of these had not been retracted despite being published for many years (Cabanac & Labbé, 2021). The screening process runs daily and results appear at the *Problematic Paper Screener*² (PPS) a public website I have been developing. This work was profiled in the *News* section of *Nature* (Van Noorden, 2021). As advised by HCÉRES-OFIS (2021), I authored 200 post peer-review comments publicly available on the PubPeer platform (Barbour & Stell, 2020). Some comments have led publishers to retract their papers; crooked authors occasionally posted vulgar replies without providing any acceptable rationale for their wrongdoings (Marcus, 2021c).

My main contribution to error detection has been to collaboratively coin the ‘tortured phrases’ concept, find thousands of such ‘tortured’ papers, and issue an Open Call for Investigation to the scientific community (Cabanac et al., 2021). The new form of plagiarism we identified involves paraphrasing to evade plagiarism detection. Fraudsters copy texts from various sources, paraphrase them, and paste them to their own articles. Paraphrasing entails the use of a thesaurus to change original words into synonyms. Established phrases such as ‘artificial intelligence’ and ‘Naïve Bayes’ (a machine learning technique named after Reverend Thomas Bayes) get synonymised as chimeric ‘counterfeit consciousness’ and ‘innocent/credulous Bayes.’ Crooked authors neither notice nor bother to correct these abnormalities that get published by flagship publishers. We view ‘tortured phrases’ as tips to focus our attention on suspect papers (Cabanac, Labbé, & Magazinov, 2022). The *News* section of *Nature* profiled our efforts to unveil this emerging form of plagiarism (Else, 2021c).

Our Cabanac–Labbé–Magazinov team detects ‘tortured’ papers by screening the entire scientific literature daily. We (re)assess the publications featuring tortured phrases and the PPS invites the scientific community at large to likewise. Post-publication evaluation reports are crowdsourced from PubPeer and we implement a snowballing approach to integrate newly found ‘tortured phrases’ for forthcoming screening batches. Research integrity sleuths and whistleblowers have joined forces, one of the most active being Elisabeth Bik whose detective work was profiled in (Shen, 2020). As of October 2021, we have flagged 2,225 problematic papers (with 1,823 ‘tortured papers’) 602 of which being commented on PubPeer by 60 individuals (Fig. 1).

We found evidence that the latest advances in text generation using deep neural networks are being diverted to produce and publish nonsensical articles. Running 140k paper abstracts through the GPT-2 detector, we found high concentrations of abstracts exhibiting a GPT

¹ <https://github.com/gcabanac/editorial-assessment>

² <https://www.irit.fr/~Guillaume.Cabanac/problematic-paper-screener>

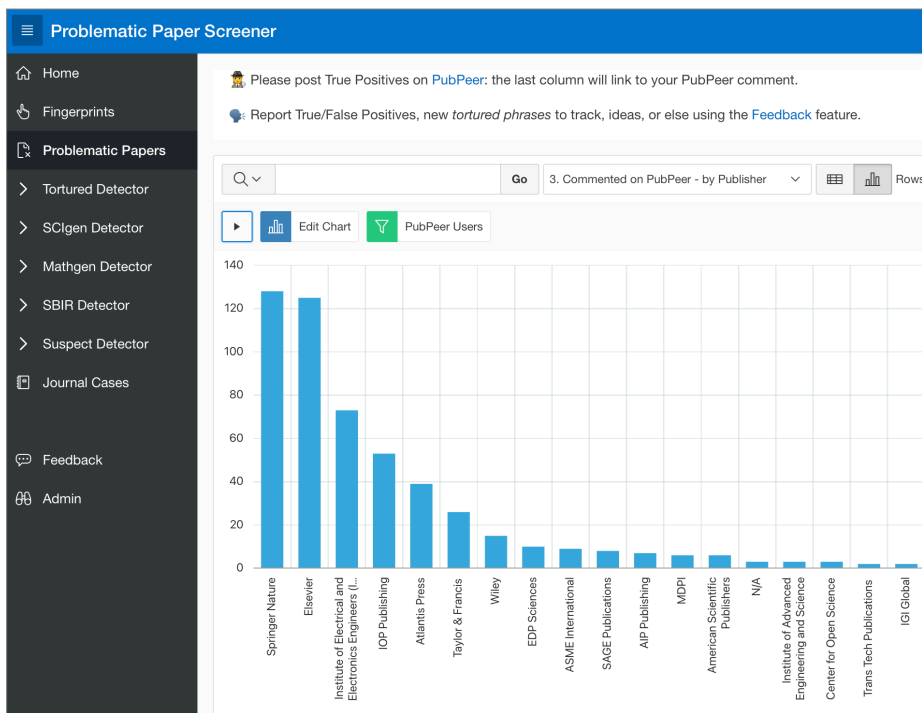


Fig. 1 Per publisher breakdown of the 602 flawed papers flagged by the PPS and commented on PubPeer. The 'fingerprints' tab lists the tracked tortured phrases (e.g., irregular timberlands) and established wordings that readers expect in the literature (e.g., random forests).

nature (i.e., 'fake' texts in GPT parlance) in the corpus of Elsevier publications of 2021 (Cabanac et al., 2021, Tab. 6). This concerning result calls for an in-depth study to determine how these were produced (machine translation, text generation, paraphrasing or else?) and assess the soundness of the associated publications.

2.2 Preventive approach to decontaminate the scientific literature

While the *curative* part of my project targets the problematic papers already published, the *preventive* part ambitions to prevent new problematic articles from integrating the scientific record in the first place. The surge of GPT-like papers we observed calls for an interdisciplinary sleuth work supported by skills in computer science, natural language processing, and artificial intelligence (Hutson, 2021; Venema, Jerde, Sweeney, & Huth, 2020). Now obsolete text generation models like GPT-2 have been used to generate scientific papers (Lang, 2019). One can only fear that the availability of enhanced models (e.g., GPT-J was trained on scientific corpora, see Wiggers, 2021) will encourage crooked authors even more.

How to anticipate the next stream of flawed papers? I envision to approach computational linguists to work on the automatic detection of unexpected phrases appearing in papers without being common in the associated journal or field. In addition, we should take advantage of emerging deceptive ways to publish nonsense being flagged

by epistemic activists like Elisabeth Bik and other users of the [PPS](#). Following [Zuckerman \(2020\)](#), I plan to keep studying the post peer-review reports and design algorithmic counterattacks to flag misconduct at the global literature scale. We should also systematically monitor publicly available data about the publishers' journal-wise output and post-peer review reports.

3 Expected contributions and outcomes

During the 5-year tenure as IUF Chair, I wish to achieve the following 5 goals.

3.1 *Designing of enhanced algorithms to comb the literature for errors*

We approached the detection of algorithmically-generated papers and tortured papers by screening the literature for certain fingerprint-queries ([Cabanac & Labbé, 2021](#)). The ongoing crowdsourcing operation supported by our [PPS](#) website reveals new tortured phrases every day. I plan to use text mining to complement this qualitative approach by digging suspect wordings from the corpora under study. An exploratory study using pointwise mutual information ([Bouma, 2009](#)) yielded promising results. Some phrases are acceptable in some fields while nonsensical in others; one needs to account for intra-field likelihood to reduce the false positive detection rate.

Some forms of error detection performed manually should benefit from automation. We have reported how misidentified cell lines³ were misused in selected biomedical publications ([Byrne et al., 2021](#); [Park et al., 2021](#)). Biological research using (or citing papers using) misidentified materials are not only a waste of time and resources but also a risk for health. The time is ripe to upscale our current digital test-tube to screen the *entire* literature, analyse flagged articles, report problematic ones via the [PPS](#), and request publishers to correct or retract them.

3.2 *Organisation of an Information Retrieval challenge*

My background in computing is in Information Retrieval (IR), which has a long experience of experimentation ([Voorhees, 2007](#)). Experimentation in IR requires a coordinating institution (e.g., NIST in the US) to specify a challenging 'search task' and invite research groups worldwide to contribute and crack the case. The algorithms produced by each participant get benchmarked against a test collection set up by organisers. The IR field advances by assessing the relative performance of the proposed methods.

The collection of problematic papers we grow daily is a valuable by-product of our error detection endeavour. I plan to set up an error-detection task at the next iteration of the *Bibliometric-enhanced Information Retrieval* I co-organised the past 7 years ([Frommholz, Cabanac, Mayr, & Verberne, 2021](#)). This will focus the attention of leading IR groups on this research integrity issue and their contributions will help to flag new misconducts. Such challenges are meeting opportunities for academics and scientists from the private sector. Elsevier is willing to support us and Philippe Terheggen, Managing Director, liaised with us shortly after Elsevier released a public statement in *Retraction Watch* ([Marcus, 2021b](#)).

³ See <https://iclac.org/databases/cross-contaminations/>.

3.3 Coordination of a distributed ‘screen and report’ task force

Our study was framed as an Open Call for Investigation to welcome the community to join forces with us ([Cabanac et al., 2021](#)). Now we are leading a crowdsourced effort to (re)assess via the [PPS](#) papers published with errors. About 1,400 articles flagged with 3+ tortured phrases are awaiting for visual inspection and commenting on PubPeer when suspected flaws get confirmed. This number increases as we add more tortured phrases to the screener, snowballing from PubPeer. We need to co-ordinate this (re)assessment effort and the way Wikipedia moderation works is inspiring. Area editors could be appointed to process the queue matching their expertise, and delegate to trusted parties part of the workload.

Science is said to be self-correcting. For this to happen, dedicated people do need to identify issues and correct the record. Authors do it when citing previous studies critically. We need to bring the critical post peer-review reports to the attention of the publishers. This we did about the SCIgen papers ([Cabanac & Labbé, 2021](#)) and publishers like ACM, IEEE, and Springer retracted the nonsensical papers after a few months. For papers not entirely nonsensical, the investigation took longer and is still hanging for some ([Byrne et al., 2021](#)). We should adopt a more systematic approach to notify publishers according to the [COPE guidelines](#) and monitor each case, reactivating our requests to investigate the cases raised. This work should benefit from and contribute to the [Retraction Watch database](#).

3.4 Knowledge and skill transfer to the publishing industry

Following the publication of ([Cabanac & Labbé, 2021](#)) integrity managers at various publishers have approached us. They typically ask to train their staff to error sleuthing and ask how to use the open source code we had released. We consider that better checks should be implemented at the publishers’ side to reduce the stream of questionable papers being published. Contracting between universities and publishers is a way to pass on knowledge via master classes or professional workshops. Some co-authors gained experience in this area and were successful at integrating their screening software into the publishers’ pipelines ([Weissgerber et al., 2021](#)). Conversely, we may learn from the experience of publishers and preprint repositories. The honorable [arXiv.org](#) screens submissions with text mining ([Ginsparg, 2014](#)) and it was found to incorrectly accept [one SCIgen paper](#) only. Tying links with the publishing industry would help to populate the collection of problematic papers required to enhance the screening process too (back to step 1 in Sect. 3.1).

3.5 Foster continuing professional education on research misconduct

Throughout the IUF tenure, I plan to follow and take part to the communications and activities of [COPE](#) (Committee on Publication Ethics) and [IRAFPA](#) (Institut de Recherche et d’Action sur la Fraude et le Plagiat) ([Bergadaà & Peixoto, 2021](#)). I intend to keep communicating about research integrity and misconduct in professional and lay venues ([Cabanac, 2021](#)). [Bergstrom and West \(2020\)](#) stress the importance of critical thinking when presented statistical analyses not to be deceived. Likewise, I believe scientists should consider research integrity as part of their continuing professional education. All researchers should take on the responsibility to decontaminate the scientific literature to pass a more sustainable scientific environment to the next generation of scholars.

4 Possible opening of IUF project towards an ERC project during IUF delegation

During the COVID19 pandemic in 2020, scientific controversies crossed the academic sphere and spilled throughout the public sphere. The “Surgisphere scandal” (Piller, 2020) represented inflated expectations for public health that relied on an information bubble stemming from a flagship journal: The Lancet. The problematic papers that had published unsupported claims sparked much hope, fuss, and confusion before being retracted eventually.

During the national lockdowns, our group of 9 researchers in France and the Netherlands worked remotely on the issue of apparent defunct self-correction mechanisms of science. We combined perspectives and skills from our diverse backgrounds in engineering (Computing, Nanobiology, and Physics) and in the social sciences (Science and Technology Studies, Sociology of Science). Our submission to the ERC Synergy call has been granted 8.3 million euros for a 5-year period starting in June 2021.

In a nutshell, Nanobubbles is concerned with how, when and why science fails to correct itself. We are looking at ways the scientific record can be corrected. Sometimes scientists make mistakes. The way scientists often think about science and non-scientists too is that those mistakes will eventually be cleared up but in practice, people who try to correct mistakes on the scientific record often experience a lot of resistance. We want to see how claims that need to be corrected circulate through scientific communities and also what happens when people contest those claims and try to get them corrected.

Nanobubbles focuses on 3 already identified bubbles (inflated expectations) in the field of nanobiology (see the executive summary [online](#)). These bubbles reflect questionable claims that the physics co-PI will reassess by replicating the experiments in his lab (a 6-month demanding research requiring staff trained in biology and physics and complex instrumentation).

During the 5-year IUF tenure, I plan to prepare an application to the ERC Advanced call. AI-powered text generation is not mature enough yet to produce sound scientific texts. The few attempts that have been discovered so far bear rhetorical errors that readers can identify. With the announced and expected advances (Romero, 2021) in this area, I believe that the next avalanche of problematic papers will stem from text generation with neural networks trained on the entire scientific corpus released in the public domain a few days ago (Else, 2021b). Research will be needed to provide editors and reviewers with software tools that detect suspect passages, checking claims against knowledge bases (like BLAST for instance), and fabricated data.

5 Possible opening of IUF project towards a project of innovation in teaching methods and dissemination throughout society

As a researcher, I design and implement information systems to process scholarly big data using text and data mining. These are built on top of databases I modelled and populated using public open data (e.g., Crossref) and subscription-based data acquired through APIs (application programming interfaces). They host 90+ million bibliographic records described with rich metadata, fulltexts, and citation links. End users browse and analyse these data using online dashboards such as the [Problematic Paper Screener](#) and the [COVID19 Preprint Monitor](#).

I approach my research and teaching activities as a continuum. The experience I gain while data harvesting, processing, and visualising myself fertilises the lectures I give and

the projects I assign students to. The 9-month project I assigned to my 3rd-year students in big data (LP GTIDM) require to process thousands subtitle files of 128 TV series to build a search and recommender system. This topic appeals to them yet I envision a future project of data and text mining of the scientific literature for them to reproduce the [Problematic Paper Screener](#).

Education to error detection in the press and social media is part of educational programmes in universities worldwide. The [syllabus](#) and book “Calling Bullshit: The art of skepticism in a data-driven world” by [Bergstrom and West \(2020\)](#) of the University of Washington is remarkable. Such training should be part of doctoral programmes (training available from the [URFIST national network](#) I am involved in as [instructor](#)), including error / misconduct / fraud detection in scientific manuscripts (when readers act as reviewers) and published papers.

Exposure to selected [PubPeer](#) posts and [Retraction Watch](#) articles could lead doctoral students, tenured faculty, and concerned citizen scientists to reflect on the self-correcting processes in science. Sometimes these work and peer-review corrects flaws before publication or publishers issue corrections and retractions (e.g., [183 retracted COVID19 studies](#)). Sometimes publishers and authors long to correct the scientific record, which jeopardises the trust the public places in science.

The unstable and ever-evolving nature of the body of knowledge has been heavily commented and sometimes criticised during the pandemic. The public needs a clearer comprehension of how research works (including error detection) and how science accumulates knowledge. Journalists have been disseminating the first results of my project on error detection to the society through radio broadcasts (e.g., [FranceInfo](#)) and press articles (e.g., [Le Monde](#), [Neue Zürcher Zeitung](#), [Frankfurter Allgemeine Zeitung](#), [Nature](#)). During my 5-year tenure as IUF Chair, I wish to keep raising the awareness on this issue while welcoming citizen scientists to contribute to the crowdsourcing of error detection we have started operating with the [Problematic Paper Screener](#).

6 Conclusion

In a nutshell, this application aims to consolidate a growing expertise on error detection in the scientific literature and develop software tools to filter out and even prevent any deceptive flaws from entering the scholarly record.

References

- Abalkina, A. (2021). Detecting a network of hijacked journals by its archive. *Scientometrics*, 126(8), 7123–7148. doi: [10.1007/s11192-021-04056-0](https://doi.org/10.1007/s11192-021-04056-0)
- Abritis, A., Marcus, A., & Oransky, I. (2020). An “alarming” and “exceptionally high” rate of COVID-19 retractions? *Accountability in Research*, 28(1), 58–59. doi: [10.1080/08989621.2020.1793675](https://doi.org/10.1080/08989621.2020.1793675)
- Baker, M. (2016). Is there a reproducibility crisis? *Nature*, 533(7604), 452–454. doi: [10.1038/533452a](https://doi.org/10.1038/533452a)
- Barbash, F. (2014, July 14). *Peer review scandal forces resignation of Taiwan cabinet minister*. Retrieved from <https://www.washingtonpost.com/news/morning-mix/wp/2014/07/14/peer-review-scandal-forces-resignation-of-taiwan-cabinet-minister/>

- Barbour, B., & Stell, B. M. (2020). PubPeer: Scientific assessment without metrics. In M. Biagioli & A. Lippman (Eds.), *Gaming the metrics: Misconduct and manipulation in academic research* (pp. 149–155). Cambridge, MA, USA: MIT Press. doi: [10.7551/mitpress/11087.003.0015](https://doi.org/10.7551/mitpress/11087.003.0015)
- Bergadaà, M., & Peixoto, P. (Eds.). (2021). *L'urgence de l'intégrité académique*. Caen: EMS Management et Société.
- Bergstrom, C. T., & West, J. D. (2020). *Calling bullshit: The art of skepticism in a data-driven world*. New York: Random House.
- Beta Writer. (2019). *Lithium-ion batteries: A machine-generated summary of current research*. Springer. doi: [10.1007/978-3-030-16800-1](https://doi.org/10.1007/978-3-030-16800-1)
- Biagioli, M., & Lippman, A. (Eds.). (2020). *Gaming the metrics: Misconduct and manipulation in academic research*. Cambridge, MA, USA: MIT Press. doi: [10.7551/mitpress/11087.001.0001](https://doi.org/10.7551/mitpress/11087.001.0001)
- Bouma, G. (2009). Normalized (pointwise) mutual information in collocation extraction. In *Proceedings of the biennial GSCL conference 2009* (pp. 31–40). Tübingen: Gunter Narr Verlag.
- Brainard, J., & You, J. (2018). What a massive database of retracted papers reveals about science publishing's "death penalty". *Science*. doi: [10.1126/science.aav8384](https://doi.org/10.1126/science.aav8384)
- Bruza, P., & Weeber, M. (2008). Literature-based discovery. In (Vol. 15). Springer. doi: [10.1007/978-3-540-68690-3](https://doi.org/10.1007/978-3-540-68690-3)
- Byrne, J. A., Park, Y., West, R. A., Capes-Davis, A., Favier, B., Cabanac, G., & Labbé, C. (2021). The thin ret(raction) line: biomedical journal responses to incorrect non-targeting nucleotide sequence reagents in human gene knockdown publications. *Scientometrics*, 126(4), 3513–3534. doi: [10.1007/s11192-021-03871-9](https://doi.org/10.1007/s11192-021-03871-9)
- Cabanac, G. (2018). What is the primordial reference for ...? — redux. *Scientometrics*, 114(2), 481–488. (Eugene GARFIELD Memorial Issue) doi: [10.1007/s11192-017-2595-4](https://doi.org/10.1007/s11192-017-2595-4)
- Cabanac, G. (2021). This digital-hygiene routine will protect your scholarship [World View: A personal take on science and society]. *Nature*, 598(7882), 541. doi: [10.1038/d41586-021-02901-z](https://doi.org/10.1038/d41586-021-02901-z)
- Cabanac, G., & Labbé, C. (2021). Prevalence of nonsensical algorithmically generated papers in the scientific literature. *Journal of the Association for Information Science and Technology*, 72(12), 1461–1476. doi: [10.1002/asi.24495](https://doi.org/10.1002/asi.24495)
- Cabanac, G., Labbé, C., & Magazinov, A. (2021, July 12). Tortured phrases: A dubious writing style emerging in science. Evidence of critical issues affecting established journals. *arXiv*. doi: [10.48550/arXiv.2107.06751](https://doi.org/10.48550/arXiv.2107.06751)
- 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*. Retrieved from <https://thebulletin.org/2022/01/bosom-peril-is-not-breast-cancer-how-weird-computer-generated-phrases-help-researchers-find-scientific-publishing-fraud/>
- Chen, X. (2019). High monetary rewards and high academic article outputs: Are Chinas research publications policy driven? *The Serials Librarian*, 77(1–2), 49–59. doi: [10.1080/0361526x.2019.1645793](https://doi.org/10.1080/0361526x.2019.1645793)
- Confronting conflict of interest [Editorial]. (2018). *Nature Medicine*, 24(11), 1629–1629. doi: [10.1038/s41591-018-0256-7](https://doi.org/10.1038/s41591-018-0256-7)
- Else, H. (2021a). Errors in genetic sequences mar hundreds of studies [News]. *Nature*. doi: [10.1038/d41586-021-02136-y](https://doi.org/10.1038/d41586-021-02136-y)

- Else, H. (2021b). Giant, free index to world's research papers released online [News]. *Nature*. doi: [10.1038/d41586-021-02895-8](https://doi.org/10.1038/d41586-021-02895-8)
- Else, H. (2021c). "Tortured phrases" give away fabricated research papers [News]. *Nature*, 596(7872), 328–329. doi: [10.1038/d41586-021-02134-0](https://doi.org/10.1038/d41586-021-02134-0)
- Else, H., & Van Noorden, R. (2021). The fight against fake-paper factories that churn out sham science. *Nature*, 591(7851), 516–519. doi: [10.1038/d41586-021-00733-5](https://doi.org/10.1038/d41586-021-00733-5)
- Faulkes, Z. (Ed.). (2021). *Stinging the predators: A collection of papers that should never have been published*. (version 18) doi: [10.6084/m9.figshare.5248264](https://doi.org/10.6084/m9.figshare.5248264)
- Frommholz, I., Cabanac, G., Mayr, P., & Verberne, S. (2021). Report on the 11th bibliometric-enhanced information retrieval workshop (BIR 2021). *SIGIR Forum*, 55(1), 1–9. doi: [10.1145/3476415.3476426](https://doi.org/10.1145/3476415.3476426)
- Gingras, Y. (2020). The transformation of the scientific paper: From knowledge to accounting unit. In M. Biagioli & A. Lippman (Eds.), *Gaming the metrics: Misconduct and manipulation in academic research* (pp. 43–55). Cambridge, MA, USA: MIT Press. doi: [10.7551/mitpress/11087.003.0004](https://doi.org/10.7551/mitpress/11087.003.0004)
- Ginsparg, P. (2014). ArXiv screens spot fake papers. *Nature*, 508(7494), 44. doi: [10.1038/508044a](https://doi.org/10.1038/508044a)
- Grudniewicz, A., Moher, D., Cobey, K. D., Bryson, G. L., Cukier, S., Allen, K., ... Lalu, M. M. (2019). Predatory journals: No definition, no defence. *Nature*, 576(7786), 210–212. doi: [10.1038/d41586-019-03759-y](https://doi.org/10.1038/d41586-019-03759-y)
- HCÉRES-OFFIS. (2021, September 27). *L'opportunité d'une discussion plus ouverte ?* Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur. Retrieved from <https://www.hceres.fr/sites/default/files/media/downloads/note-ofis-cofis-20210927.pdf> (Note d'expression de l'OFFIS et de son conseil)
- Herzog, C., Hook, D., & Konkiel, S. (2020). Dimensions: Bringing down barriers between scientometricians and data. *Quantitative Science Studies*, 1(1), 387–395. doi: [10.1162/qss_a_00020](https://doi.org/10.1162/qss_a_00020)
- Hutson, M. (2021). Robo-writers: the rise and risks of language-generating AI [News feature]. *Nature*, 591(7848), 22–25. doi: [10.1038/d41586-021-00530-0](https://doi.org/10.1038/d41586-021-00530-0)
- Labbé, C., Cabanac, G., West, R. A., Gautier, T., Favier, B., & Byrne, J. A. (2020). Flagging incorrect nucleotide sequence reagents in biomedical papers: To what extent does the leading publication format impede automatic error detection? *Scientometrics*, 124(2), 1139–1156. doi: [10.1007/s11192-020-03463-z](https://doi.org/10.1007/s11192-020-03463-z)
- Labbé, C., Grima, N., Gautier, T., Favier, B., & Byrne, J. A. (2019). Semi-automated fact-checking of nucleotide sequence reagents in biomedical research publications: The Seek & Blastn tool. *PLOS ONE*, 14(3), e0213266. doi: [10.1371/journal.pone.0213266](https://doi.org/10.1371/journal.pone.0213266)
- Lang, F. (2019, October 28). *OpenAI's GPT2 now writes scientific paper abstracts*. Interesting Engineering. Retrieved from <https://interestingengineering.com/openai-gpt2-now-writes-scientific-paper-abstracts>
- Lewison, G., & Sullivan, R. (2015). Conflicts of interest statements on biomedical papers. *Scientometrics*, 102(3), 2151–2159. doi: [10.1007/s11192-014-1507-0](https://doi.org/10.1007/s11192-014-1507-0)
- Lin, S. (2013). Why serious academic fraud occurs in China. *Learned Publishing*, 26(1), 24–27. doi: [10.1087/20130105](https://doi.org/10.1087/20130105)
- Maisonobe, M., Jégou, L., & Cabanac, G. (2018). Peripheral forces: The growing impact of second-tier cities is narrowing the gap in research production. *Nature*, 563(7729), S18–S19. (Nature Index 2018 Science Cities supplement) doi: [10.1038/d41586-018-07210-6](https://doi.org/10.1038/d41586-018-07210-6)

- Marcus, A. (2021a). Elsevier says “integrity and rigor” of peer review for 400 papers fell “beneath the high standards expected”. *Retraction Watch*. Retrieved from <https://retractionwatch.com/?p=122620>
- Marcus, A. (2021b). ‘Preprints are works in progress’: The tale of a disappearing COVID-19 paper. *Retraction Watch*. Retrieved from <https://retractionwatch.com/?p=122145>
- Marcus, A. (2021c). ‘Tortured phrases’, lost in translation: Sleuths find even more problems at journal that just flagged 400 papers. *Retraction Watch*. Retrieved from <https://retractionwatch.com/?p=122688>
- Merton, R. K. (1942). Science and technology in a democratic order. *Journal of Legal and Political Sociology*, 1(1), 115–126. doi: [2027/mdp.39015008014428](https://doi.org/10.2307/39015008014428)
- Oikonomidi, T., Boutron, I., Pierre, O., Cabanac, G., Ravaud, P., & the COVID-19 NMA Consortium. (2020). Changes in evidence for studies assessing interventions for COVID-19 reported in preprints: Meta-research study. *BMC Medicine*, 18(402). doi: [10.1186/s12916-020-01880-8](https://doi.org/10.1186/s12916-020-01880-8)
- Oransky, I. (2021a). Guest editor says journal will retract dozens of inappropriate papers after his email was hacked. *Retraction Watch*. Retrieved from <https://retractionwatch.com/?p=122972>
- Oransky, I. (2021b). Introducing two sites that claim to sell authorships on scientific papers. *Retraction Watch*. Retrieved from <https://retractionwatch.com/?p=123031>
- Oransky, I. (2021c). Springer Nature slaps more than 400 papers with expressions of concern all at once. *Retraction Watch*. Retrieved from <https://retractionwatch.com/?p=123181>
- Park, Y., West, R. A., Pathmendra, P., Favier, B., Stoeger, T., Capes-Davis, A., ... Byrne, J. A. (2021, July 31). Human gene function publications that describe wrongly identified nucleotide sequence reagents are unacceptably frequent within the genetics literature. *bioRxiv*. doi: [10.1101/2021.07.29.453321](https://doi.org/10.1101/2021.07.29.453321)
- Phillips, N. (2017). Nature’s 10: Ten people who mattered this year — Jennifer Byrne: Error sleuth. *Nature*, 522(7685), 321. doi: [10.1038/d41586-017-07763-y](https://doi.org/10.1038/d41586-017-07763-y)
- Piller, C. (2020). Who’s to blame? These three scientists are at the heart of the Surgisphere COVID-19 scandal. *Science*. doi: [10.1126/science.abd2252](https://doi.org/10.1126/science.abd2252)
- Pupovac, V. (2021). The frequency of plagiarism identified by text-matching software in scientific articles: A systematic review and meta-analysis. *Scientometrics*, 126(11), 8981–9003. doi: [10.1007/s11192-021-04140-5](https://doi.org/10.1007/s11192-021-04140-5)
- Romero, A. (2021, September 11). *GPT-4 will have 100 trillion parameters—500x the size of GPT-3*. Towards Data Science. Retrieved from <https://towardsdatascience.com/gpt-4-will-have-100-trillion-parameters-500x-the-size-of-gpt-3-582b98d82253>
- Sharma, K. (2021). Team size and retracted citations reveal the patterns of retractions from 1981 to 2020. *Scientometrics*, 126(10), 8363–8374. doi: [10.1007/s11192-021-04125-4](https://doi.org/10.1007/s11192-021-04125-4)
- Shen, H. (2020). Meet this super-spotter of duplicated images in science papers. *Nature*, 581(7807), 132–136. doi: [10.1038/d41586-020-01363-z](https://doi.org/10.1038/d41586-020-01363-z)
- Soete, L., Schneegans, S., Eröcal, D., Angathevar, B., & Rasiah, R. (2015). A world in search of an effective growth strategy. In S. Schneegans (Ed.), *UNESCO Science Report: Towards 2030* (pp. 20–55). Paris. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000235406>
- Teixeira da Silva, J. A. (2021). Abuse of ORCID’s weaknesses by authors who use paper mills. *Scientometrics*, 126(7), 6119–6125. doi: [10.1007/s11192-021-03996-x](https://doi.org/10.1007/s11192-021-03996-x)
- Van Noorden, R. (2021). Hundreds of gibberish papers still lurk in the scientific literature [News]. *Nature*, 594(7862), 160–161. doi: [10.1038/d41586-021-01436-7](https://doi.org/10.1038/d41586-021-01436-7)

-
- Venema, L., Jerde, T., Sweeney, Y., & Huth, J. (2020). New chapter in intelligence writing [Editorial]. *Nature Machine Intelligence*, 2(8), 419. doi: [10.1038/s42256-020-0223-0](https://doi.org/10.1038/s42256-020-0223-0)
- Voorhees, E. M. (2007). TREC: Continuing information retrieval's tradition of experimentation. *Communications of the ACM*, 50(11), 51–54. doi: [10.1145/1297797.1297822](https://doi.org/10.1145/1297797.1297822)
- Weissgerber, T., Riedel, N., Kilicoglu, H., Labbé, C., Eckmann, P., ter Riet, G., ... Bandrowski, A. (2021). Automated screening of COVID-19 preprints: Can we help authors to improve transparency and reproducibility? *Nature Medicine*, 27(1), 6–7. doi: [10.1038/s41591-020-01203-7](https://doi.org/10.1038/s41591-020-01203-7)
- Wiggers, K. (2021, June 9). *EleutherAI claims new NLP model approaches GPT-3-level performance*. Venturebeat. Retrieved from <https://venturebeat.com/2021/06/09/eleutherai>
- Zuckerman, H. (2020). Is “the time ripe” for quantitative research on misconduct in science? *Quantitative Science Studies*, 1(3), 945–958. doi: [10.1162/qss_a_00065](https://doi.org/10.1162/qss_a_00065)