U.S. HOUSE OF REPRESENTATIVES
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY
SUBCOMMITTEE ON INVESTIGATIONS AND OVERSIGHT

HEARING CHARTER

Paper Mills and Research Misconduct: Facing the Challenges of Scientific Publishing

Wednesday, July 20, 2022
10:00 a.m. EDT – 12:00 p.m. EDT
Zoom

PURPOSE

The purpose of this hearing is to discuss the current and future challenges in securing scientific literature from fraudulent academic papers. The hearing will examine field-specific and industry-wide strategies for identifying fraud, the increasing number of fraudulent papers produced and sold by paper mills, and the impact of new technologies such as AI on both the perpetration and the detection of research misconduct. Members and witnesses will discuss the successes of the largely volunteer post-publication review community, the challenges that community has faced, and the strategies publishers themselves are developing to combat research misconduct.

WITNESSES

- **Dr. Jennifer Byrne**, Director, Biobanking, New South Wales Health Pathology; Professor of Molecular Oncology, University of Sydney
- **Mr. Chris Graf,** Research Integrity Director, Springer Nature; Chair, Governance Board, STM Association Integrity Hub
- **Dr. Brandon Stell,** Neuroscientist, French National Centre for Scientific Research; President and Co-Founder, The PubPeer Foundation

OVERARCHING QUESTIONS

- What is the scope and what are the sources of misconduct in academic publishing?
- What tools and methodologies exist to detect research misconduct both prior to and following publication?
- How are these tools being deployed and who is responsible for deploying them?
- How will automation help or hinder the fight against research misconduct in the scientific literature?

BACKGROUND

What is Research Misconduct?

In December of 2000, the Office of Science and Technology Policy (OSTP) defined research misconduct as "fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results," harmonizing conflicting definitions across the US Federal scientific enterprise.[1] Fabrication refers to creating data out of whole cloth, falsification is manipulating results to give a desired outcome, and plagiarism is the appropriation of another's work. Any of these alone is sufficient to identify an action as research misconduct, and all fall under the broad umbrella of academic fraud.

The OSTP definition, and the subsequent investigative procedures, specifically covered research conducted with Federal funds. However, since research misconduct renders the conclusions of research invalid, no papers which are the product of misconduct should be present in the scientific literature, whether or not they were supported with public funds. Ideally such papers would be identified early and rejected from the publishing process, either during peer review or during evaluation by research integrity professionals employed at the journals themselves. Unfortunately, researchers studying academic fraud identify numerous published papers which contain evidence of research misconduct every year.

Why commit research misconduct?

Individual research misconduct is when a researcher or researchers commit falsification, fabrication, or plagiarism to make their work publishable. Techniques for committing this fraud can vary by field. For instance, in the life sciences key data often takes the form of images such as cellular assays or western blots, which are relatively easy to digitally alter to better reflect a desired result. Other fraud might include altering datasets by adding or removing data to better make it demonstrate a desired or significant result.

In academia there is a concept colloquially referred to as "publish or perish," i.e., that a certain quantity of publications is required to advance in a career. In the U.S., the number of publications, the number of citations on those publications, and the impact factor of the journal that prints a researcher's work are common metrics to assess a researcher's quality as a scientist and their worthiness for recruitment or prestigious teaching positions.[2] Conducting research is just one among many responsibilities of most academics, and even when performed well, publication-worthy results are not guaranteed. As such, there is an incentive to supplement output by using research misconduct as a shortcut. Recognizing this problem, some organizations have taken steps to standardize alternative metrics for assessing research quality.[3] For instance, the San Francisco Declaration on Research Assessment pushed for a greater focus on the quality of research rather than the number or location of publications.[4]

---

[1] https://www.govinfo.gov/content/pkg/FR-2000-12-06/pdf/00-30852.pdf
[2] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3999612/
[3] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2126010/?tool=pubmed
[4] https://sfdora.org/read/

Incentives for research misconduct can be even stronger in other countries. In Russia, the 2012 May decrees included a push for certain universities to increase their international ranking, in part by increasing their number of publications.[5] In China, physicians were required to publish research articles, often in English language journals, to gain promotions or even be permitted to perform certain surgeries, regardless of whether they possessed the time or resources to conduct such research.[6] China has taken steps to remove explicit cash incentives for publishing low quality papers,[7] but there are still strong incentives to publish a paper, any paper, to move to the next level in a career.[8] So long as the academic community, both domestic and international, values publication quantity over quality, the incentive to commit research misconduct will remain.

Paper Mills

The challenge of scaling research misconduct oversight has grown along with the incidence of "paper mills." Paper mills are organizations, often based in Russia or China, which sell scientific papers, authorship positions on papers, or other fraudulent academic products.[9] They may disguise themselves as offering translation services for non-native English-speaking researchers.[10] Some will create a fake email account and entirely take over the correspondence with journals on behalf of the customer. One research integrity expert told the Committee that it is a strong indicator of a paper mill when a non-native English-speaking researcher is too responsive or too willing to share data or other information to move along their paper submission. While some use sloppy and unprofessional formatting, others are careful to match the submission standards of the journals to which they send their fraudulent papers.

Paper mills have a multitude of strategies to source the fraudulent papers they sell and/or attempt to publish for a price. One technique is to contact researchers with offers to pay to add fake co-authors to legitimate research. Another strategy is to take over the editing process by bribing editors to turn a blind eye to fraudulent papers in the submission queue or by purchasing editorial control from unscrupulous editors.[11] They have also been known to hijack entire journals either by assuming the digital identity of the journal,[12] or by using identity fraud to pose as a guest editor of a journal to guarantee that their own papers are accepted.[13] Paper mills may also create their own papers en masse through ghostwriting, templates, or the application of automated tools that will be discussed in more detail below.

No complete census of these organizations exists, and the scope of their impact is difficult to determine. But anecdotally, independent researchers who investigated one Russian paper mill were able to identify hundreds of papers it had produced and published, leading to the retraction

[5] https://arxiv.org/ftp/arxiv/papers/2112/2112.13322.pdf
[6] https://www.nature.com/articles/d41586-021-00733-5
[7] https://www.nature.com/articles/d41586-020-00574-8
[8] https://publicationethics.org/resources/forum-discussions/publishing-manipulation-paper-mills
[9] https://www.science.org/content/article/russian-website-peddles-authorships-linked-reputable-journals
[10] https://publicationethics.org/resources/forum-discussions/publishing-manipulation-paper-mills
[11] https://retractionwatch.com/2018/05/30/want-to-earn-10k-per-month-join-the-journals-mafia/
[12] https://retractionwatch.com/2022/05/29/want-to-know-whether-that-journal-is-scamming-you-introducing-the-retraction-watch-hijacked-journal-checker/
[13] https://retractionwatch.com/2021/06/18/galling-journal-scammed-by-guest-editor-impersonator/

of 30 papers from one journal alone.[14] Another paper mill based in China, dubbed "the tadpole paper mill" for the shape of the fraudulent western blot images it produced, has been linked to more than 400 papers published primarily across six journals.[15]

ENTITIES COMBATTING MISCONDUCT IN ACADEMIC PUBLISHING

Publishers

Academic publishers have a vested interest in protecting the reputations of their journals by detecting fraudulent articles prior to publication. Peer review, in which subject matter experts are recruited to examine a paper within their discipline for worthiness, is the classic quality control method of academic publishing. However, it is unreasonable to expect subject matter experts to also excel at forensic examinations of images, data, and author backgrounds. As such, checks for research misconduct must occur during other phases of the publishing process, an effort which requires its own dedicated resources.

Publishers have several avenues for cooperation to help detect and mitigate misconduct in the papers they publish. One example is the Committee on Publication Ethics (COPE), which was formed in 1997 to coordinate on misconduct cases.[16] COPE issues guidance to publishers, provides example protocols for handling research misconduct, and creates flowcharts for specific fraud types such as image manipulation[17] or systemic fraud.[18] COPE began discussing paper mills in earnest in the fall of 2020[19] and, in a partnership with the International Association of Scientific, Technical and Medical Publishers (STM) released a paper analyzing paper mills in depth in June of 2022.[20]

STM has also begun developing a collaboration platform for the 140 publishers that comprise their membership. This platform is called the STM Integrity Hub and Chris Graf chairs the governance board. The first major component of this project will be to develop a system that will detect simultaneous submissions to multiple journals – a strong indicator of paper mill activity – while protecting the intellectual property of the publishers and the privacy of paper authors.[21] Future efforts will aim to facilitate the development and exchange of research integrity best practices and the creation of additional tools to detect research misconduct.

Volunteer Post-Publication review community

Despite the efforts of publishers, many fraudulent papers are detected following publication by a primarily volunteer post-publication review community. Many members of this community started as academics and then altered or expanded their trajectory after a serendipitous encounter with research misconduct. For instance, Dr. Elisabeth Bik was an academic microbiologist by

[14] https://retractionwatch.com/2022/07/05/our-deepest-apology-journal-retracts-30-likely-paper-mill-articles-after-investigation-published-by-retraction-watch/
[15] https://scienceintegritydigest.com/2020/02/21/the-tadpole-paper-mill/
[16] https://publicationethics.org/about/history
[17] https://publicationethics.org/files/image-manipulation-published-article-cope-flowchart.pdf
[18] https://publicationethics.org/files/publication-process-manipulation-cope-flowchart.pdf
[19] https://publicationethics.org/resources/forum-discussions/publishing-manipulation-paper-mills
[20] https://publicationethics.org/files/paper-mills-cope-stm-research-report.pdf
[21] https://www.stm-assoc.org/stm-integrity-hub/

training who now makes a living through her expertise at detecting image manipulation and her services as a research integrity consultant.[22] Others, such as Dr. Jennifer Byrne, are academics who have broadened their area of study while maintaining a traditional professorship at a research university. While she eventually received a small grant from the Office of Research integrity, Dr. Byrne's work scanning papers for wrongly identified nucleotide sequencing reagents was initially undertaken without funding.[23] Still other volunteers examine raw numerical data for signs of manipulation using statistical tests.[24]

These volunteers, many of them anonymous, conduct ad-hoc reviews based on their respective interests and skills. There is no overarching organization independent of journals themselves to fund fraud detection or target it to where it is most needed within the scientific literature. The quantity of papers detected by volunteers lacking a systemic approach suggests they are in a target rich environment, i.e., the full scope of fraud within the scientific literature is larger than is currently known. Unfortunately, some volunteers who identify research misconduct report harassment and lawsuits from authors they have called out,[25] and others stay anonymous in part to avoid receiving that kind of reaction.[26] Other volunteers reported giving up on notifying journals of their findings after repeatedly being stonewalled while clearly flawed papers were allowed to stand unchallenged.

To combat fraud as effectively as possible, these volunteers have coalesced around several websites for communication and coordination, including Twitter.[27] The blog Retraction Watch is also a valuable resource for the community as a provider of the latest news on fraud detection, a place to post findings that might not fit in traditional journals, and as a database of all journal retractions.[28] Another website, PubPeer.com – founded by Dr. Brandon Stell – provides a place for anyone to comment on concerns on scientific articles published in any journal. The website provides a space for authors to discuss potential problems in their work directly with the people who detected those problems.[29]

Federal Government

A finding of research misconduct could end a scientist's career, so Federal science agencies follow a structured process for the investigation and adjudication of allegations of scientific misconduct.[30] This process applies strictly to research funded by Federal agencies.

One Federal leader for detecting research misconduct is the Office of Research Integrity (ORI) housed within the Department of Health and Human Services. ORI's stated mission is to oversee all Public Health Service (PHS) research integrity issues.[31] It performs this mission in part by conducting investigations into cases of research misconduct using PHS funding and by providing

[22] https://scienceintegritydigest.com/about/
[23] https://www.nature.com/articles/d41586-021-02136-y
[24] https://www.science.org/content/article/meet-data-thugs-out-expose-shoddy-and-questionable-research
[25] https://www.the-scientist.com/news-opinion/elisabeth-bik-faces-legal-action-after-criticizing-studies-68831
[26] https://forbetterscience.com/2019/07/30/help-with-another-not-on-pubpeer-yet/
[27] *Ibid.*
[28] https://retractionwatch.com/the-center-for-scientific-integrity/
[29] https://pubpeer.com/static/about
[30] https://www.govinfo.gov/content/pkg/FR-2000-12-06/pdf/00-30852.pdf
[31] https://ori.hhs.gov/policies-ori-mission

small grants – between $50K and $150K – for research, development, and demonstration of research integrity projects.[32] ORI's investigations are thorough, in keeping with the OSTP guidelines from the definition of research misconduct, but they are also relatively few, averaging just 4 case summaries published each year since 2016.[33]

Though it does not offer research grants, the NSF Office of the Inspector General performs a similar function to ORI by conducting research misconduct investigations when there are credible allegations of misconduct in NSF-funded research. The OIG has tools to detect plagiarism. It also relies on tips from the public or reporting from the volunteer review community, including through Retraction Watch and PubPeer, to bring fabrication or falsification cases to their attention for evaluation.

Automation

Automated tools for translating or modifying text so that it can pass through a traditional plagiarism detector are being used to help to disguise plagiarism in scientific papers.[34] In 2005, graduate students at MIT developed a tool called SCIgen which could produce a nonsensical but well formatted computer science research paper on demand. Researchers using the tool managed to get several papers published in a variety of journals over the years as a pointed statement on journal vetting processes.[35] Another robo-writer tool is designed specifically to produce Small Business Innovative Research (SBIR) proposals. At this time, the results are generally jargon-heavy to disguise their fundamentally nonsensical nature but could easily be detected by an attentive human with relevant expertise. Recent developments have brought incredible strides in AI-powered language models. Tools such as GPT-3 show an incredible versatility with a range of writing formats, and robo-written scientific material will only grow more sophisticated along with these tools.[36]

However, while automation may be enabling more efficiencies for committing research misconduct, it also offers new opportunities for detection and mitigation. Automated tools can enable publishers and volunteer reviewers to evaluate papers more quickly and thoroughly than would be possible for a human working unaided. Note that even advocates stress that humans should always make the final determination of research misconduct when automated tools are used to assist with detection.

Detecting plagiarism: The detection of direct plagiarism has long been the focus of research misconduct tools. Millions of students have had their papers checked by Turnitin.com since the tool's creation in 1998, and publishers commonly employ a variant called iThenticate or other similar services to check for plagiarism in submitted articles.[37]

Researchers have discovered a way to detect the use of robo-writers through specific "tortured phrases" where pairs of phrases are technically synonyms, but a human would clearly see a

[32] https://ori.hhs.gov/blog/fy-2022-grant-opportunity-forecasts
[33] https://ori.hhs.gov/content/case_summary
[34] https://www.nature.com/articles/d41586-021-02134-0
[35] https://pdos.csail.mit.edu/archive/scigen/#people
[36] https://www.technologyreview.com/2020/07/20/1005454/openai-machine-learning-language-generator-gpt-3-nlp/
[37] https://www.turnitin.com/products/ithenticate

distinction. For instance, there is a clear difference between "breast cancer" and "bosom peril." Once a tortured phrase such as bosom peril is identified it can serve as an easily detectable marker of a plagiarized paper.[38] A study searching for tortured phrases found almost 8,000 in the published literature, many in reputable journals.[39]

Detecting fabrication: OpenAI, which creates the GPT series of AI-powered language models, has also produced a detector for identifying text created by the GPT-2 version of the tool. When researchers applied that tool to Elsevier paper abstracts, it flagged hundreds of papers for containing synthetic text likely produced by GPT-2.[40]

Detecting falsification: The tadpole paper mill was found in part because hundreds of fraudulent papers all used similar images with an eponymous tadpole-like structure[41]. Multiple companies, such as Image Twin[42] and Proofig,[43] are developing tools that will automatically detect manipulation or duplication of images across papers. Those tools have shown early promise, but they have not yet been externally evaluated or scaled up to the level needed to handle even the hundreds of fraudulent images known to exist.

Another tool called Seek & Blastn was created to detect highly field-specific fraud. It scans papers and extracts gene identifiers and nucleotide sequences from the text and flags potential incorrect usages of reagents and sequences for humans to review .[44] Using this tool, a team of researchers was able to reduce a population of 12,000 papers down to 3,400 with potential problems and ultimately identified 712 flawed papers.[45] This tool is limited in scope, as it cannot recognize non-human genomes, but it demonstrates that automated tools can provide valuable assistance even when the fraud in question requires significant scientific expertise to recognize.

[38] https://arxiv.org/abs/2107.06751
[39] https://dbrech.irit.fr/pls/apex/f?p=9999:24::IR_years
[40] https://thebulletin.org/2022/01/bosom-peril-is-not-breast-cancer-how-weird-computer-generated-phrases-help-researchers-find-scientific-publishing-fraud/?utm_source=Twitter&utm_medium=SocialMedia&utm_campaign=TwitterPost01132022&utm_content=DisruptiveTech_TorturedPhrasesPointToFraud_01132022
[41] https://scienceintegritydigest.com/2020/02/21/the-tadpole-paper-mill/
[42] https://imagetwin.ai/
[43] https://www.proofig.com/
[44] https://www.protocols.io/view/seek-amp-blastn-standard-operating-procedure-q26g7b2k1lwz/v1
[45] https://pubmed.ncbi.nlm.nih.gov/35022248/