Straive™

Connecting the dots

# Uncovering the Paper Mill Industry: A Comprehensive Approach to Addressing Paper Mills in Scholarly Publishing

WHITEPAPER

# 1.Introduction

Plagiarism, data fabrication, and falsification are the most common forms of scientific misconduct that lead to the retraction of academic works. As technology advances, so do the ways in which scientific fraud and misconduct can occur. One such example is the rise of paper mills. These are commercial enterprises that produce large volumes of academic papers for sale, often without any regard for academic standards or ethical guidelines. The proliferation of paper mills has led to a surge in plagiarism, the dissemination of false or misleading information, and the devaluation of academic publishing as a whole.

Academic paper mills have become a growing concern in the scholarly publishing industry, causing significant harm to the integrity of academic research and scholarship. In recent years, there has been a growing recognition of the negative impact that paper mills can have on the academic community. Institutions and publishers have increasingly implemented measures to detect and prevent the use of papers produced by these mills, but much work remains to be done. It is essential to raise awareness of the dangers of paper mills and the importance of taking action to address this issue.

In this paper, we will provide a comprehensive overview about paper mills and their impact on scholarly publishing. We will examine the ways in which paper mills operate, the types of services they offer, and the negative consequences they can have on academic research and publishing. Additionally, we will explore the importance of addressing this issue and the measures that can be taken to combat paper mills effectively.

# 2.Understanding Paper Mills:

The emergence of paper mills is a growing cause for concern for academic publishers because of the widespread retraction of papers that follows when the practice is uncovered. Up to 2% of papers submitted to journals are suspected of being fake, according to a recent study conducted jointly by the Council on Publication Ethics (COPE) and the International Association of Scientific, Technical, and Medical Publishers (STM) (COPE & STM, 2022).

## What are Paper Mills and why would researchers use them

Paper Mills are unethical businesses that benefit from authors who pay for their services: creating academic papers frequently based on falsified or modified data and submitting them to scholarly publications. As such, it violates every cardinal rule of responsible scientific research and publication. In addition, this is a systemic process that threatens to derail the accumulation of information upon which scientific advancement depends.

Paper mills come in many varieties, each offering its own unique set of services (ghost writing, experiments, data falsification, complete article sales, etc.) for various rates.

Restorative Neurology and Neuroscience's Editor-in-Chief, Professor Bernhard Sabel, recently gave a talk at the APE 2022 in which he discussed the results of an investigation into whether or not his publication had fallen prey to paper mills[1]. They identified 23 (15.3%) "likely" fake papers (FPs). They then used a random sampling method to screen 3500 articles from 35 basic and clinical neuroscience journals, where 376 "suspects" (10.7%) were uncovered. Extrapolating this figure to PubMed publications, they projected 450.000 medical FPs and 1.4 million FPs for all 14 million science/technology publications.

Identified **23 (15.3%)** "likely" fake papers (FPs).

**376** "suspects" **(10.7%)** uncovered.

**2022**

Researchers, students, and low-ranking institutions who are under pressure to boost their statistics or publication totals may often resort to unscrupulous practices such as paper mills. The motivation is either monetary or one of prestige. For instance, a researcher may decide to pay for paper mill papers rather than perform the research themselves if they need to publish a particular number of papers to be considered for a promotion or raise.

From assisting academics with their experiments to producing full papers using fabricated data, paper mills provide a wide range of services.

## How do Paper Mills Operate and Profit:

Many paper mills capitalize on the pressure on academics to publish in order to stay competitive. They provide rapid publication chances for a price, usually between several hundred and several thousand dollars per manuscript, but without adequate peer review or editorial oversight. Some paper mills also provide "reputation management" services, in which they write favorable articles about a person to raise their academic standing.

The sole motivation for offering this service is financial gain. Authorships on ready-to-submit articles command significant fees from researchers. Potential customers include researchers looking for a shortcut to publishing in high-impact, peer-reviewed international journals with little to no actual research input. Certain paper mills also have fully functioning labs where experiments are conducted and real data and images are generated. Some authors purchase these data as well to use in various experiments.

Manuscripts created for authors who desire to submit them to journals with an impact factor of more than 3 can cost as much as 30,000 EUR. Paper mills are a significant industry, valued at about 2 billion euros[2].

# The SME Role in identifying Paper Mills

Most of the time, paper mill submissions are not detected until after the work has been published. This is due to the fact that identifying and recalling fabricated contributions requires a significant number of data points in many systems.

On the grounds of fabricated research[3], the Royal Society of Chemistry retracted 68 publications from RSC Advances, RSC Medicinal Chemistry, and Food and Function in January 2021. RSC, in collaboration with experts, launched a yearlong investigation that uncovered multiple articles with evidence of possible fraud. These seemingly credible papers shared "extremely identical structures or templates" despite having different authors. Many instances of image duplication and alteration were also identified.

These manuscripts first seem legitimate independently; common themes and shared characteristics aren't revealed until editors compare numerous articles written by various researchers who have nothing in common. While there exist a variety of sophisticated text-based plagiarism detection tools, tools that can identify instances of data-based plagiarism are still rather uncommon.

Subject Matter Experts (SMEs) are specialists who are knowledgeable about the literature that has already been published in their area of expertise. They are able to spot when a paper claims to have new findings that are not supported by previous research.

The editors of a journal may seek raw data if they have doubts about a manuscript. Yet, verifying the veracity of data is not simple, especially if the examination of data files requires specialized tools. This procedure may be time-consuming and costly. It may also be difficult to tell if the data has been manipulated unless you are an expert in the field.

Some signs that SMEs can check for to spot paper mills are:

- Unusual or repetitive author names
- Poor quality of writing or language
- Lack of references or citations
- Unfamiliar or unknown institutions or affiliations
- Inconsistent or incomplete data
- Unusual or suspicious submission patterns
- High volume of submissions from a single author or institution
- Lack of response to editorial queries or requests for revisions
- Unusual or suspicious payment arrangements
- Use of fake or manipulated images or data

However, it is important to note that these signs may not always indicate the presence of a paper mill, but they can serve as red flags for further investigation.

## 3.Consequences of Paper Mills

The use of papers produced by paper mills compromises the credibility of academic research and publishing because of the high rates of plagiarism and poor quality of the content they produce. While the exact impact of these mills is difficult to quantify, there are several negative consequences that can be observed.

Paper mills undermine the peer-review process, an essential part of the scholarly publishing process, whereby experts in a given field assess the quality and validity of a research paper before it is published. Paper mills circumvent this process entirely by providing pre-written papers, which can be easily purchased and submitted for academic credit. This means that students and researchers are no longer required to undergo the rigorous peer-review process, which can result in the publication of low-quality research.

The authenticity of a researcher's work is severely compromised when they resort to using pre-written papers or plagiarized content. This might lead to inaccurate or misleading conclusions, which can have serious ramifications, especially in fields like medicine where research findings can affect patient care. The usage of pre-written papers might also inhibit researchers from actively participating in the research process and developing their own critical thinking skills, both of which are essential to the academic pursuit.

Moreover, the increasing frequency of retracted papers originating from paper mills is posing a problem for the research community.

A cross-sectional study published in the British Medical Journal found that paper mill retractions accounted for 21.8% of all retractions in 2021 (Fig 1).

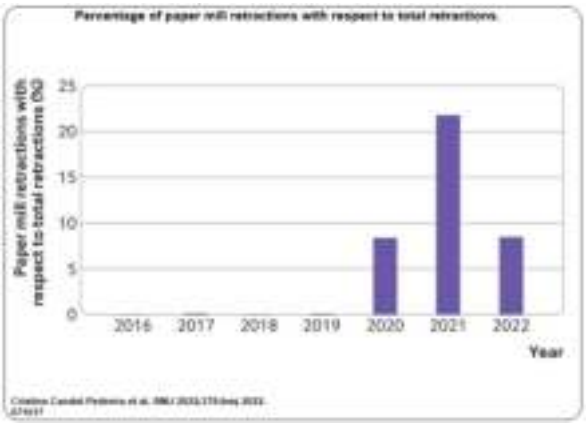



Fig 1: Percentage of paper mill retractions with respect to total retractions
Source: theBMJ

The study identified 1182 retracted paper mill papers published between 2004 and 2022, with 93.8% of these papers having received at least one citation since publication. The study also found that almost all authors of paper mill retractions came from Chinese institutions, followed by far fewer authors from Indian institutions (Fig 2).

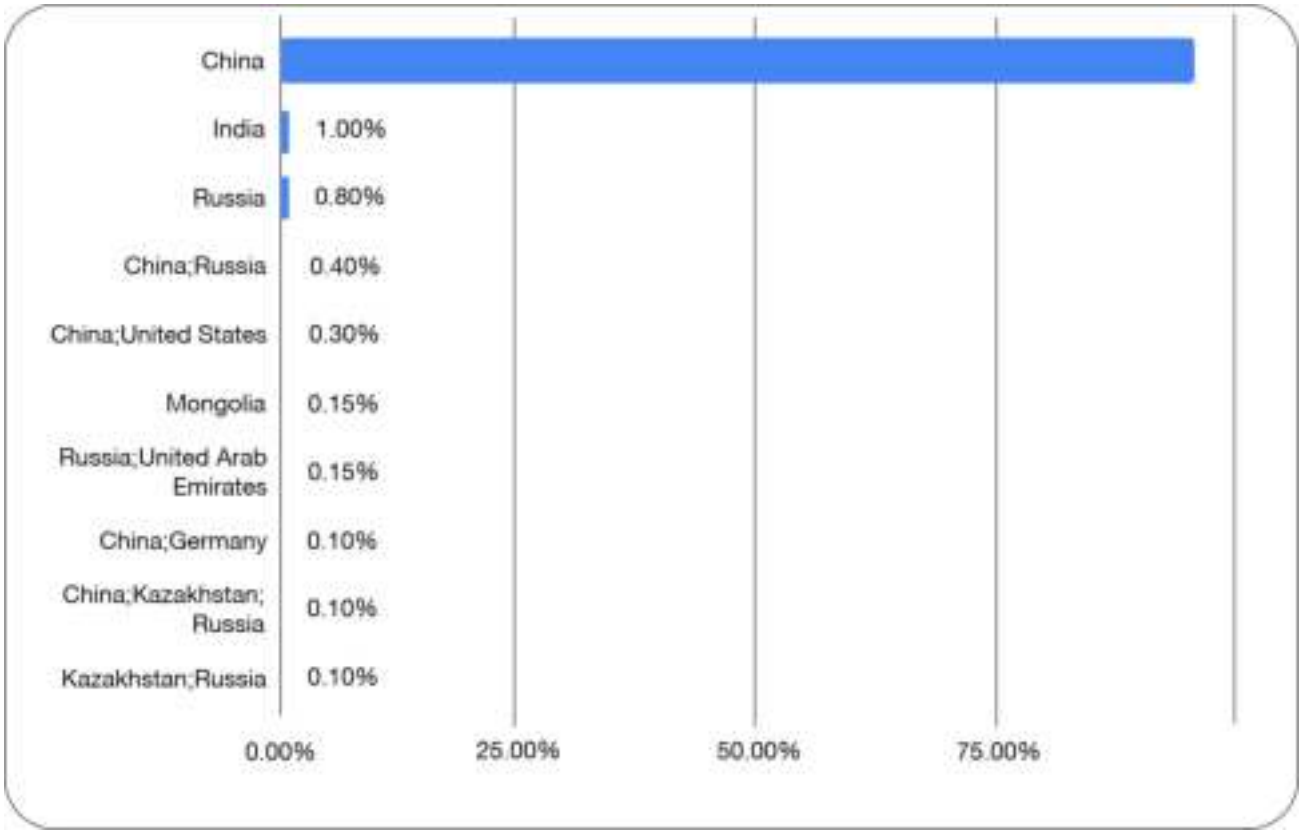



Fig 1: Percentage of paper mill retractions based on country

Paper mills also have significant financial implications. These companies profit by selling pre-written papers, which can be purchased at a low cost. This means that they are able to undercut legitimate publishers, who invest significant resources into the peer-review process and the production of high-quality research. The financial implications of paper mills are compounded by the fact that they often operate in countries with liberal copyright laws, making it difficult for publishers to take legal action against them.

As such, it is crucial that publishers work to combat the rise of paper mills and ensure that the integrity of the scholarly publishing process is maintained. This may include implementing new measures to detect and prevent plagiarism, increasing transparency and accountability in the publishing process, and educating students and other stakeholders on the importance of academic integrity.

# 4.Existing efforts to combat Paper Mills

The proliferation of paper mills poses a serious threat to the scholarly publishing sector as well as to the integrity of academic research. While several initiatives have been taken to combat them, their effectiveness is limited by the complexity of the issue. The lack of coordination between different stakeholders, the global reach of paper mills, and the rapid growth ofthese services make it challenging to combat them effectively. Nonetheless, continued efforts are essential to protect the integrity of scholarly publishing and academic research.

## Initiatives to combat Paper Mills

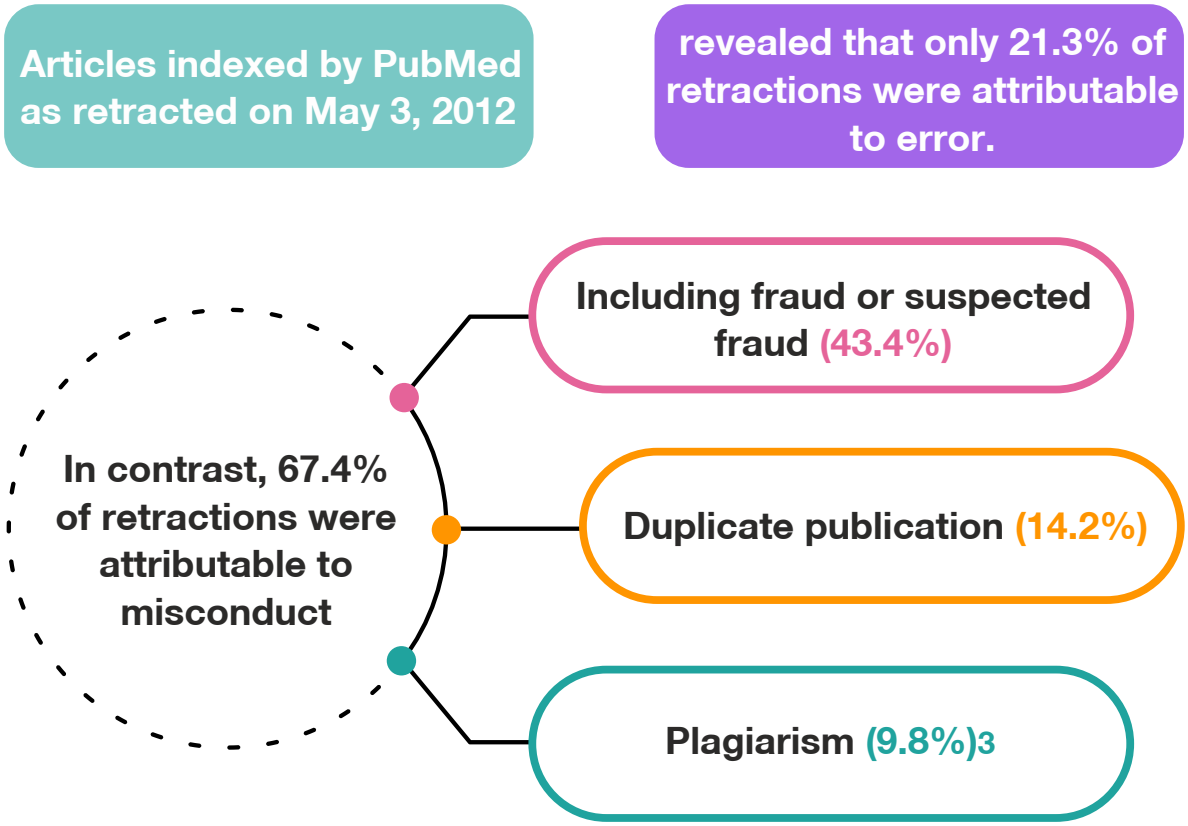
Several initiatives have been put in place to combat Paper Mills. One notable initiative is the use of plagiarism detection software that can help detect instances of plagiarism and duplication of content. Many academic institutions also require researchers to use anti-plagiarism tools such as iThenticate and Crossref Similarity Check before submitting their work for publication.

Another initiative is the use of digital object identifiers (DOIs) to uniquely identify scholarly publications. DOIs help to ensure the authenticity and integrity of scholarly publications, making it easier to track instances of plagiarism and fraudulent publication.

In addition, academic institutions and publishers have also established guidelines and standards for the publication and dissemination of scholarly research. For example, the Committee on Publication Ethics (COPE) provides guidelines for ethical publication practices, while the International Committee of Medical Journal Editors (ICMJE) has established guidelines for the conduct, reporting, editing, and publication of scholarly research in medical journals.

## Effectiveness & limitations of these initiatives

While these initiatives have been successful in detecting and preventing instances of plagiarism and fraudulent publication, they have not completely eradicated the problem of paper mills. The complex nature of the issue, the global reach of these services, and the lack of coordination between different stakeholders make it difficult to implement comprehensive solutions. A detailed review of 2,047 biomedical and life-science research




Identification software, for example, is only effective in identifying papers that have already been submitted or published. This means that new papers produced by paper mills may not be detected. Legal action can be effective in some cases, but it is expensive and time-consuming. Moreover, paper mills often operate in countries where legal action is difficult to pursue.

Public awareness campaigns can help to reduce the demand for paper mills, but they have a limited impact on those who are already using these services. Moreover, the lack of standardization in academic policies across different institutions makes it challenging to enforce policies related to academic integrity.

While several initiatives have been put in place to combat Paper Mills, there is a need for more proactive measures to prevent the occurrence of plagiarism and fraudulent publication, including greater collaboration between academic institutions and publishers, and more rigorous enforcement of ethical publication standards.

## 5.Best Practices to Address Paper Mills

Addressing paper mills in the scholarly publishing industry requires a comprehensive approach that involves collaboration among stakeholders, the development of policies and guidelines, education and training for researchers and editors, and the use of emerging technologies. By implementing these best practices, the scientific community can ensure that high-quality research is published and the integrity of academic publishing is maintained.

## Collaboration among Stakeholders

Collaboration among stakeholders, including publishers, editors, researchers, and academic institutions, is essential to address paper mills. It is crucial to identify the root causes of paper mills and understand the motivation behind them. For example, researchers may resort to paper mills due to the pressure to publish their research or the lack of resources to conduct high-quality research. Academic institutions may also be incentivized to publish research articles to secure funding or attract students. By working together, stakeholders can develop solutions that address these underlying issues and prevent paper mills from proliferating. Initiatives such as the STM Integrity Hub can serve as a valuable platform for facilitating this collaboration and sharing best practices and resources to prevent and detect fraudulent research.

## Development of Comprehensive Policies and Guidelines


To prevent paper mills from operating, publishers and academic institutions should develop comprehensive policies and guidelines that address ethical publishing practices. These policies should include guidelines for authors, editors, and reviewers to ensure that all parties adhere to ethical publishing practices. Publishers should also implement rigorous review processes that detect fraudulent research articles and ensure that they are not published. Academic institutions should also provide support to researchers, such as access to resources and funding, to help them conduct high-quality research.

## Education and Training for Researchers and Editors

Education and training for researchers and editors are essential in preventing and identifying paper mills in scholarly publishing. Publishers can provide resources to help researchers identify and report paper mills, including guidelines for detecting plagiarism and fraud. They can also offer training to editors on best practices for identifying fraudulent research papers. Academic institutions can provide training on research ethics, plagiarism, and citation practices to help researchers avoid publishing in paper mills. Such training can help researchers better understand the importance of publishing in reputable journals and the potential consequences of engaging in fraudulent publishing practices.

## 6.Emerging Technologies and Their Potential to Combat Paper Mills

The issue of paper mills and unethical scholarly publishing practices has been a longstanding problem within the academic community. However, emerging technologies are now offering potential solutions to combat these issues and ensure the integrity of academic research.

Artificial intelligence (AI) is playing a significant role in the fight against paper mills and unethical publishing practices. AI can be used to analyze large volumes of data and identify patterns that may be indicative of fraudulent activity. This can help publishers and researchers quickly identify and remove fraudulent papers, reducing the amount of time and resources spent on peer review and other verification processes. Additionally, AI can be used to create more accurate and effective plagiarism detection tools, making it easier to identify instances of plagiarism and prevent it from occurring in the first place.

Another technology that is being used to counter paper mills is machine learning. By analyzing large datasets of academic papers, machine learning algorithms can identify patterns and trends that may indicate fraudulent or unethical practices. This can help researchers and publishers to identify potential issues before they become widespread, and take action to prevent them from occurring.

Open access publishing models are also helping to combat paper mills and unethical publishing practices by making academic research more widely available to the public. By removing barriers to access, open access publishing can help to promote transparency and accountability within the academic community, while also increasing the visibility and impact of research.

By leveraging the power of AI, machine learning, and open access publishing models, researchers and publishers can work together to ensure the integrity of academic research, and promote transparency and accountability within the academic community. While there are certainly challenges to implementing these technologies, including issues related to data privacy and security, the benefits of using these tools are clear. As publishers and researchers continue to explore the potential of these technologies, we can expect to see even more innovative solutions emerge in the years to come.v

## 7.Future Directions for tackling Paper Mills

The issue of paper mills continues to plague the academic community. Tackling paper mills and promoting academic integrity requires a multi-faceted approach that addresses the underlying issues that contribute to the problem. Future directions should focus on a multi-pronged approach that addresses both the supply and demand sides of the problem.

Efforts should be made to identify and shut down paper mills through legal action, enforcement of copyright laws, and increased collaboration between academic institutions and publishers. In addition, there needs to be a greater emphasis on education and awareness of academic integrity and the consequences of academic misconduct. This can be accomplished through training programs for researchers, as well as public campaigns that highlight the importance of honesty and integrity in scholarly publishing.

Reactive measures could involve stronger enforcement of existing academic misconduct policies, including more severe consequences for those found to be using paper mills. This could include revoking degrees, suspension or expulsion from educational institutions, and legal action. Academic journals and publishers should also implement stricter guidelines for the submission and review of research papers to ensure their validity and authenticity.

Another potential direction is to promote greater transparency and accountability in scholarly publishing. This can be achieved through the use of open-access platforms, pre-registration of research studies, and other measures that encourage greater sharing of data and methods. By making research more transparent and accessible, it may be easier to identify cases of academic misconduct and hold individuals or institutions accountable for their actions.

Another promising approach is to continue developing and implementing automated tools that can detect plagiarism and other forms of academic misconduct. These tools can analyze large amounts of data and identify suspicious patterns or language usage that may indicate misconduct. In addition, machine learning algorithms can be used to improve the accuracy of these tools over time.

One area of ongoing research is the development of machine learning and natural language processing algorithms to detect and identify fraudulent articles. This is a complex task, as paper mills use various techniques to hide their fraudulent activities, such as using plagiarized text, manipulating images, and using fake peer reviewers. However, there have been some promising developments in this area, such as the use of deep learning techniques to identify linguistic patterns that are common in fraudulent articles.

Another area of research is the development of new metrics and indicators to identify papers that may be fraudulent. Traditional metrics such as citation counts and journal impact factor are not always reliable indicators of quality, as paper mills may use tactics such as self-citation to artificially boost their metrics. Some new metrics that have been proposed include the detection of irregular citation patterns, the use of sentiment analysis to identify overly positive language, and the analysis of co-authorship networks to identify potential collaboration patterns.

Collaboration between educational institutions, academic publishers, and funding agencies is crucial in tackling this issue comprehensively. By continuing to develop new technologies, metrics, and policies, and by working together to implement these strategies, the academic community can take meaningful steps toward ensuring integrity and quality in scholarly publishing.

## 8.The Straive Advantage

At Straive, we recognize the importance of research integrity and publishing ethics in the scientific community. We believe that providing solutions that combine technology, process, and data can help achieve an optimal balance between protecting against bad actors, improving author education, and ensuring that appropriate manuscripts are not rejected due to false alarms, all while respecting publishing times and budgets.


To achieve this, we employ a combination of human expertise and AI-ML tools to provide accurate and reliable research support services. Our team of experienced professionals is dedicated to ensuring that research is conducted ethically and transparently. We work closely with our clients to understand their needs and tailor our services accordingly, providing customized solutions that meet their specific requirements.

We believe that transparency is essential in research, and our solutions are designed to promote transparency at every stage of the research process. Our technology and process solutions help identify potential ethical issues early on, allowing our clients to address them proactively.

We understand that time and budget constraints are always a consideration when it comes to research, and our solutions are designed to work within these constraints. We strive to provide efficient and cost-effective services that deliver results. Visit <Straive RI page> to know more.

# 9.Conclusion

The prevalence of paper mills in scholarly publishing is a significant concern for the academic community. The increasing number of papers being published in predatory journals is alarming, and the consequences can be damagi to scientific progress, public trust, and the integrity of the research community. Addressing paper mills in scholarly publishing requires a comprehensive approach that include collaboration between publishers, researchers, funding agencies, and academic institutions.

The efforts made so far in combating paper mills in scholarly publishing are commendable. However, we must continue to work towards eliminating these unethical practices by increasing awareness, developing robust policies, and enforcing stricter regulations. Researchers, academic institutions, and funding agencies must work together to promote transparency, integrity, and ethical standards in scientific research and publication.The efforts made so far in combating paper mills in scholarly publishing are commendable. However, we must continue to work towards eliminating these unethical practices by increasing awareness, developing

robust policies, and enforcing stricter regulations. Researchers, academic institutions, and funding agencies must work together to promote transparency, integrity, and ethical standards in scientific research and publication.

It is crucial to ensure that only high-quality, peer-reviewed research is published, and it is up to all stakeholders to maintain the integrity of the scholarly publishing industry. Therefore, we must remain vigilant in our efforts to combat paper mills in scholarly publishing and continue to promote a culture of transparency, rigor, and ethical conduct. Only by working together can we ensure that scientific research continues to make a positive impact on society.

## About Straive

Straive is a market-leading content technology enterprise that provides data services, subject matter expertise (SME), and technology solutions to multiple domains, such as research content, eLearning/EdTech, and data/information providers. With a client base scoping 30 countries worldwide, Straive's multi-geographical resource pool is strategically located in seven countries - the Philippines, India, the United States, Nicaragua, Vietnam, the United Kingdom, and Singapore, where the company is headquartered.

www.straive.com
straiveteam@straive.com



# References:

Nash, J. (2022, May 9). Paper Mills—The Dark Side of the Academic Publishing Industry
Tackling paper mills and bogus research: Some things are better done together | Advancing Discovery | Springer Nature. (2022, April 28)
Fang, F. C., Steen, R. G., & Casadevall, A. (2012, October). Misconduct accounts for the majority of retracted scientific publications. Proceedings of the National Academy of Sciences, 109(42), 17028–17033
RSC Advances Editorial: retraction of falsified manuscripts. (2021). RSC Advances, 11(7), 4194–4195Duine, M. (2022, December 16). Summary report APE 2022 The future of the permanent record.
Information Services & Use, 42(3–4), 261–274
Paper mills research. (2022, June 20)
Candal-Pedreira et al., "Retracted Papers Originating from Paper Mills: Cross Sectional Study." BMJ, e071517