Today's session

In this session we will

- Describe the problem of papermills
- Outline the issues
- Discuss what to look out for
- Share resources and advice for dealing with them
- Consider potential future initiatives
- Invite feedback and questions

Papermills: Overview

- Papermills manufacture manuscripts and submit them to journals on behalf of researchers for a fee.

- Submissions from papermills are not the work of the named author and they contain fraudulent content.

- Detection of articles from papermills is difficult. Current detection tools may not show up problems.

- Papers emerging from paper mills are more easily detected at scale, many are being discovered by internet sleuths who are alerting publishers and editors.

- An increasing number of these submissions are being discovered. Dealing with papermills is costly, detrimental to reputations of all involved, and makes the literature less trustworthy.

- Many of these submissions are to medical journals – they are spreading misinformation which could cost lives

Papermills: why do they exist?

- The promotion and graduation rules set by funders and institutions are driving the use of papermills and until these change, the papermill business will continue to boom.

- Authors pay for these articles to be created and published because they need publications for career progression or graduation.

- Many of the named authors who purchase these papers are doctors who may have poor English, and have full caseloads, but are still expected to publish in a journal with an Impact Factor to progress or graduate.

Papermills: scale of the problem

- So far over 1300 articles have been discovered in the literature, with more being discovered all the time.

- It is believed that there are 1000s more articles still undiscovered in the literature and that they are still being submitted.

- 26% of discovered articles have been retracted or have had expressions of concern added. Many of the remainder are still being investigated.

- The main countries where these articles originate are China, Iran and Russia, although there are other countries also involved.

- The papermills are aware of attempts to discover their papers and are able to change the way they operate to avoid detection. So detection tools need to be constantly upgraded to take account of the way they are changing.

Papermills: further reading

- Potential paper mills and what to do about them
  - https://publicationethics.org/publishers-perspective-paper-mills

- The fight against fake-paper factories that churn out sham science
  - https://www.nature.com/articles/d41586-021-00733-5 -

- Introducing two sites that claim to sell authorships on scientific papers
  - https://retractionwatch.com/2021/09/07/introducing-two-sites-that-claim-to-sell-authorships-on-scientific-papers/

Experience from small journals

Papermills: present, but not much visible to journals in small scientific communities
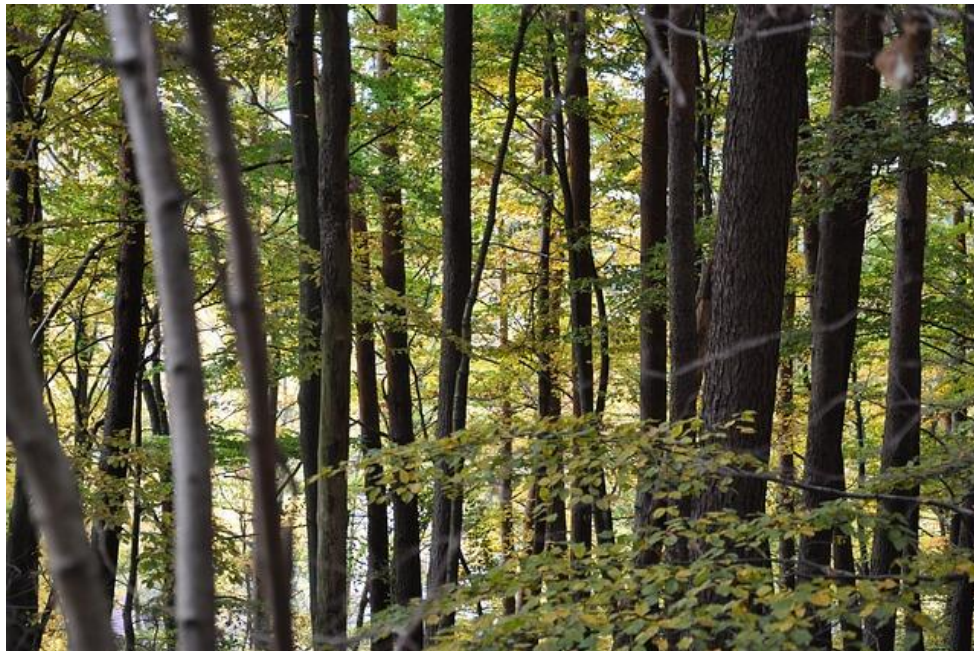
Examples of papermills in the „scientific periphery"

Exclusive: Russian site says it has brokered authorships for more than 10,000 researchers

Персонал

Журналы

Клиенты

Отзывы

Офис

ISI Web of SCIENCE

Scopus

THOMSON REUTERS

ВАК России и многие другие

Базы журналов

A company in Russia hawks its wares

Retraction Watch

> We sell publications of finished articles in Scopus and Web of Science magazines (articles written and accepted in journals; sold in parts or in whole).

Examples of papermills in the „scientific periphery"

Iran

ISI Collaborative Article Printing or ISI Collaborative Paper Printing is one way to help researchers who are having difficulty doing their research; Also, researchers who do not have sufficient expertise to write an article and, for example, are not fluent in English or specialized editing; Or people who do not have the facilities (space, tools, workshop, etc.) to do the research they want completely and sufficiently; Or those who do not have the necessary knowledge to collect data and do not know the research method completely, and many other researchers with various issues. This group of people can use the ISI collaborative article printing method and contribute to the writing of the article as possible.

Examples of papermills in the „scientific periphery"

Latvia

Science Publisher Company
Publication Support Services for Academic, Scientific Manuscripts and Papers

Home   About ⌄   Co

Co-authorship

Co-authorship in Scopus, Web of Science research papers

Co-authorship provides an opportunity to become a co-author of an existing article that has been accepted for publication in **Scopus, Web of Science (SCIE, SSCI, AHCI, ESCI)** indexing journals.

Root causes of papermills in the „scientific periphery"

Scientific publications in 2016

publicationethics.org
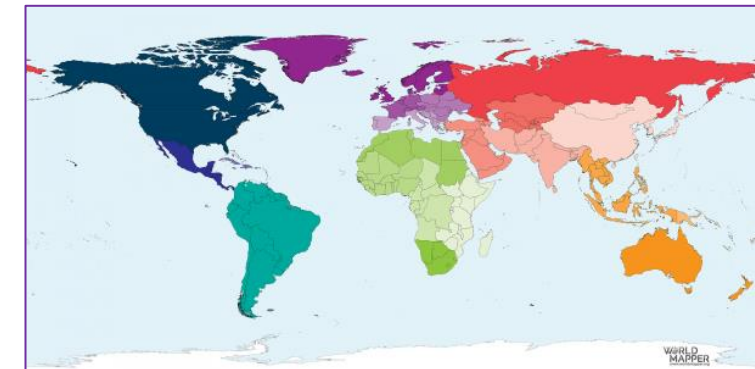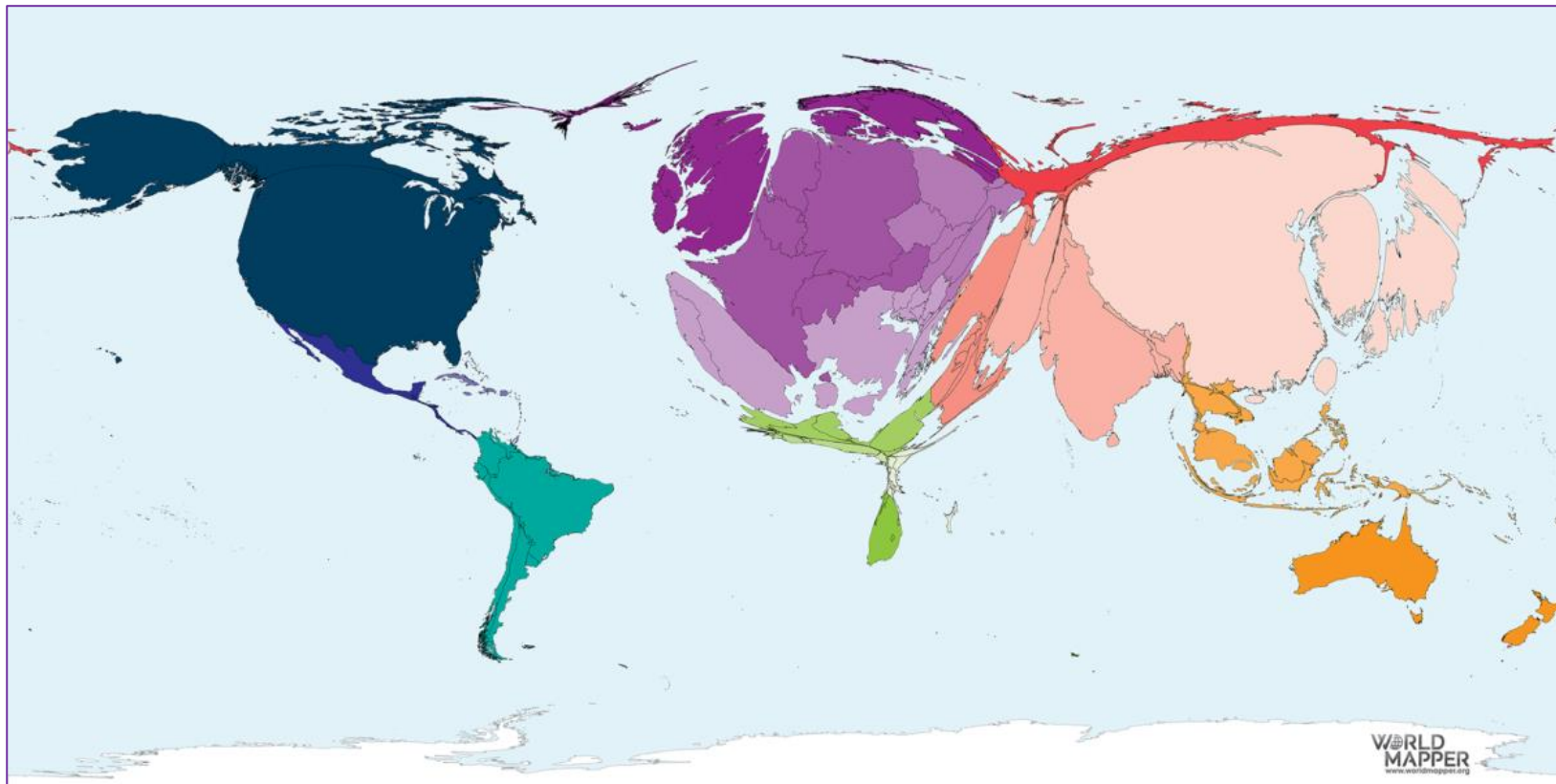
Root causes of papermills in the „scientific periphery"

Vicious circle for small scientific journals

Author pool	No. and quality of manuscripts	**LOW**
+	↓	+
Reviewer's pool	Review process	**INADEQUATE**
+	↓	+
Finances	Technical resources	**LOW**
+	↓	+
Language	English	**IMPERFECT**
↓	↓	↓
Product	Journal	**LOW QUALITY**

Poor visibility

Marušić A, Marušić M. Croat Med J 1999;40:508-514.
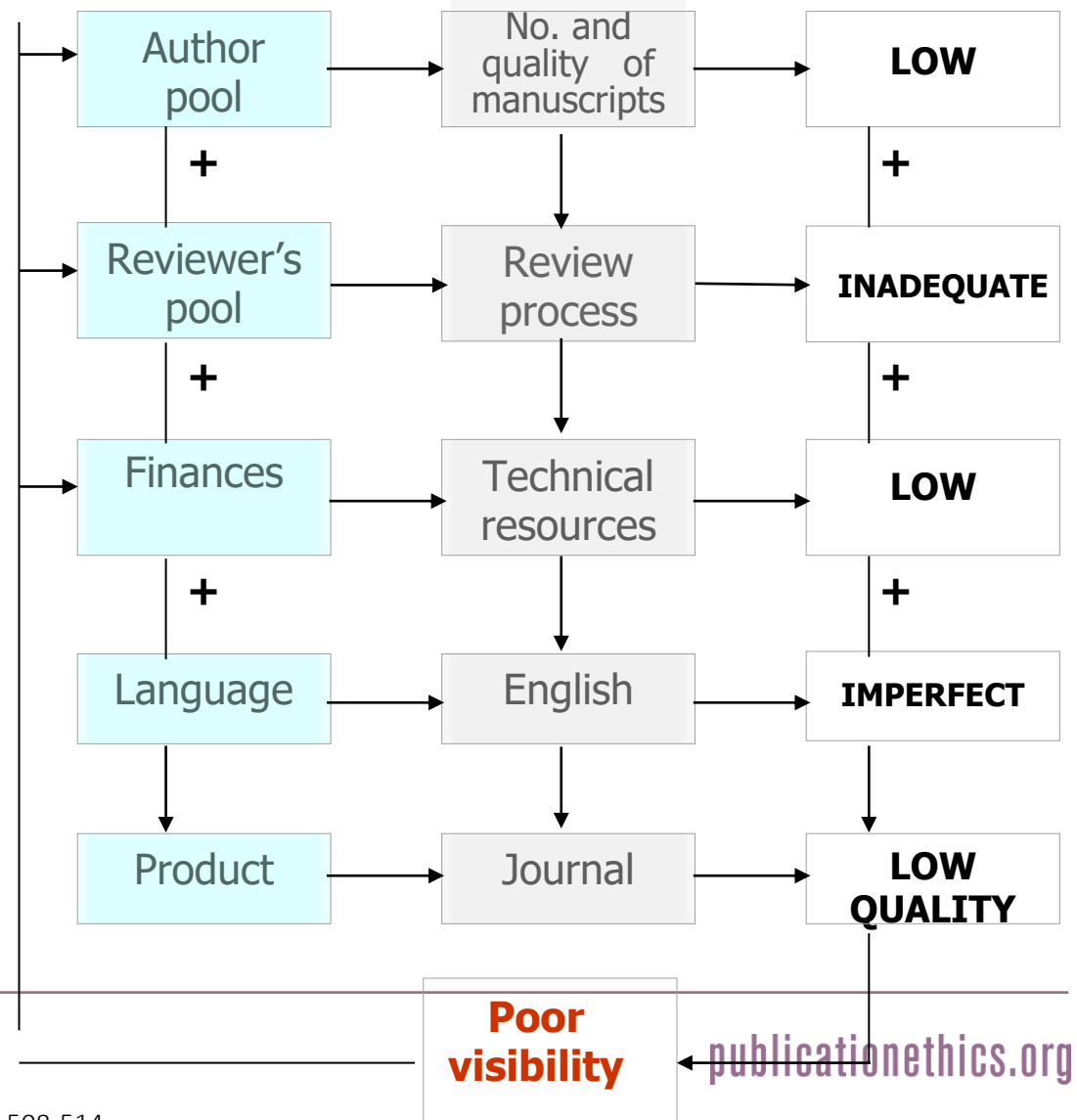
publicationethics.org

How to detect and manage paper mills

Dr Jigisha Patel

publicationethics.org

'Paper mills' = Manipulation of the publication process

COPE definition of the phenomenon is

Systematic manipulation of the publication process

By an individual or a group of individuals

Use dishonest or fraudulent practices to:

    prevent independent peer review
    sell authorship
    publish fabricated or plagiarised research

The goal is to influence the publication record and/or achieve financial gain

Who runs paper mills?

Individual or groups of researchers

Third party agencies

Openly operating paper mill businesses

How paper mills work

Step 1: Generate manuscripts

Fabricate them

Standard reporting formats

Data types that are easy to manipulate – images of Western blots

Get researchers to bring them to you

Special Issues (with guest editors)

Steal them

Peer reviewers

And many more!

How paper mills work

Step 2: Guarantee publication

Take over the peer review process to ensure that manuscripts are accepted

Offer editors bribes to accept manuscripts

Target vulnerable journals

Any many more!

How paper mills work

Step 3: Collect payment

Once a manuscript is accepted, offer authorship for sale

Pay on publication

Authorship changes and peer review manipulation don't have to be features of 'paper mills'

How are paper mills detected?

Field experts

Detect fabricated data in published articles

Elizabeth Bik & Jennifer Byrne

And other 'whistle blowers'

Publishers

Suspicious patterns of behaviour in publication workflows and databases

Editors

'Red flags' in manuscript content

publicationethics.org

How to investigate a suspicion

Journal level

Content

Institutions

Scope, plagiarism, declarations etc

Ethics & consent for research on humans

Manuscript title

Author 1, Author 2, Author 3

Institution 1, Institution 2

Text Text Text Text Text
Text Text Text Text Text
Text Text Text Text Text
Text Text Text Text Text
Text Text Text Text Text
Text Text Text Text Text
Text Text Text Text Text
Text Text Text Text Text
Text Text Text Text Text
Text Text Text Text Text

Authorship

Data

Images

Peer review

How to investigate a suspicion

Publisher level

Patterns of behaviour

Submission systems
User databases (across different journals)
Submission numbers and other features

Peer review
Emails characteristics (non-institutional, shared etc)
Frequent peer reviewers
Peer reviewer turnaround times

(COPE guidelines: How to recognise potential manipulation of the peer review process)

How to investigate a suspicion

New COPE guidelines and flowcharts

Coming soon

More details on what to look for

In submission systems

Peer review

Manuscript content

Managing a paper mills

New COPE guidelines and flowcharts

Coming soon

**Suspicious features categorized into levels of confidence that they represent
a paper mill**

Low	Medium	High
Features alone do not undermine the manuscript or article and may be legitimate behaviour	Features alone do not undermine the manuscript or article and require further information or clarification	Features alone undermine confidence in the manuscript or article enough to justify further action

Managing a papermills

PROMOTING INTEGRITY IN SCHOLARLY
RESEARCH AND ITS PUBLICATION

Common questions

Should you ask for raw data?

Should you involve institutions?

Can you publish an expression of concern or a retraction?
        If so, which?

Raw data

Do you need the raw data to decide what to do?

BEFORE you ask for data, decide:

What is the minimum information you need to verify the data?

How will you review the raw data?

Do you have the expertise to judge yourself?
Do you need expert reviewers?

COPE Seminar 2021 | How to detect and manage paper mills | Jigisha Patel

publicationethics.org

Institutions

When and why should you involve institutions?

To help with your investigation?

Institutions cannot investigate manipulation of your systems

Can help with specific questions

Did the authors do the research at that institution
Confirm the existence of data (to identify fabricated manuscripts)
Investigate concerns about research ethics in human studies

Consider informing institutions that you are investigating a paper mill

C O P E

Final action

Retract or publish an expression of concern?

COPE retraction guidelines

For example:

- Peer review manipulation
- Unethical research
- Plagiarism
- Unreliable findings

Expression of concern

- When findings are less clear-cut but concerns remain
- Subjective judgement on how far the reported research is undermined

*New COPE guidelines under development

How to detect and manage paper mills

Summary

Be aware of how manipulation of the publication process works

Know what to look for in individual manuscripts, publication workflows and databases

Be clear about when to ask for raw data and when to contact institutions

Final action depends on your confidence in the integrity of your content (remember COPE resources)

COPE resources

COPE retraction guidelines:

https://publicationethics.org/retraction-guidelines

COPE Systematic manipulation of the publication process flowcharts:

https://publicationethics.org/resources/flowcharts/systematic-manipulation-publication-process

How to recognise potential manipulation of the peer review process:

https://publicationethics.org/resources/flowcharts/how-recognise-potential-manipulation-peer-review-proces

COPE forum discussion on Expressions of Concern:

https://publicationethics.org/resources/forum-discussions/expressions-of-concern

STM at a glance: member organization of publishers

- STM supports its members in advancing trusted research worldwide.

- Over 140 members, in over 20 countries around the world.

- Includes academic and professional publishers, learned societies, and university presses; includes established players as well as start-ups.

- STM covers all scholarly disciplines.

- All members together publish 66% of journal articles and 10,000's of monographs and reference works.

- STM **works together** to serve science and society by developing standards and technology to ensure research is of **high quality, trustworthy** and easy to access.

Science became

- **Larger**
- **More international**
- **More complex**

Nature.com

Global scientific output doubles every nine years

07 May 2014 | 16:46 BST | Posted by Richard Van Noorden | Category: Policy, Publishing

Rank	Region, country, or economy	2008	2018
-	World	1,755,850	2,555,959
1	China	249,049	528,263
2	United States	393,979	422,808
3	India	48,998	135,788
4	Germany	91,904	104,396
5	Japan	108,241	98,793
6	United Kingdom	91,358	97,681
7	Russia	31,798	81,579
8	Italy	56,157	71,240
9	South Korea	44,094	66,376
10	France	66,460	66,352
11	Brazil	35,490	60,148
12	Canada	53,296	59,968
13	Spain	44,191	54,537
14	Australia	37,174	53,610
15	Iran	17,034	48,306
-	EU	528,938	622,125

https://ncses.nsf.gov/

Science became

- **Larger**
- **More international**
- **More complex**

And so did scientific fraud and misconduct!

Nature.com

Global scientific output doubles every nine years

Rank	Region, country, or economy	2008	2018
-	World	1,755,850	2,555,959
1	China	249,049	528,263
2	United States	393,979	422,808
3	India	48,998	135,788
4	Germany	91,904	104,396
5	Japan	108,241	98,793
6	United Kingdom	91,358	97,681
7	Russia	31,798	81,579
8	Italy	56,157	71,240
9	South Korea	44,094	66,376
10	France	66,460	66,352
11	Brazil	35,490	60,148
12	Canada	53,296	59,968
13	Spain	44,191	54,537
14	Australia	37,174	53,610
15	Iran	17,034	48,306
-	EU	528,938	622,125

https://ncses.nsf.gov/

Research Integrity is part & parcel of *Advancing Trusted Research*

- Quality, ethics and integrity are key values of our industry, and safeguarding them is a key role and value-add for scholarly publishers

- Technology is taking centre stage in this space – both as an enabler for 'good' (detecting and protecting) but also for 'bad', e.g. fabricated articles from paper mills.

- Publishers are gearing up to rise to the challenge:
  - Research Integrity officers & groups
  - Development of tools
  - Participation in cross-publisher Working Group and Task Forces

- STM has spearheaded cross-publisher collaboration via STEC working groups on e.g. Image Alteration & Duplication Detection

STM Advancing trusted research

Recommendations for handling image integrity issues

STM Working Group on Image Alteration and Duplication Detection[1]

- Image alteration and duplication often indications of papermills
- Best-practice recommendations that outline a structured approach to support editors and others applying image integrity screening
- Currently in consultation phase (until end of October)
- https://www.stm-assoc.org/standards-technology/working-group-on-image-alterations-and-duplications/
- https://osf.io/xp58v/

Collaboration is key

- Collaboration is <u>necessary</u>:
  - Speed and agility are critical in a tech 'arms race' (Paper Mills won't wait)
  - Better outcomes and greater ROI when sharing (IT) resources and expertise
  - More and more problems need a joint solution, where no publisher can solve it on their own – especially very high barrier for smaller & mid-sized publishers

- Collaboration is <u>realistic</u>:
  - Publishers are willing to share algorithms, ideas and expertise
  - Publishers are willing to 'pool' content when (strict) criteria for security & confidentiality are met
  - Publishers are willing to work with trusted third parties for Research Integrity services (e.g. iThenticate)

- Collaboration is <u>difficult</u>:
  - Interplay of legal, policy, workflow and technology issues
  - We are mindful of anti-trust regulation and don't want competition with our own members
  - Poor interoperability between publishers makes it inefficient to collaborate and exchange algorithms / content.
  - Re-inventing the wheel: Paper Mills, Image Alteration, Plagiarism Detection, Authorship Validation, etc.
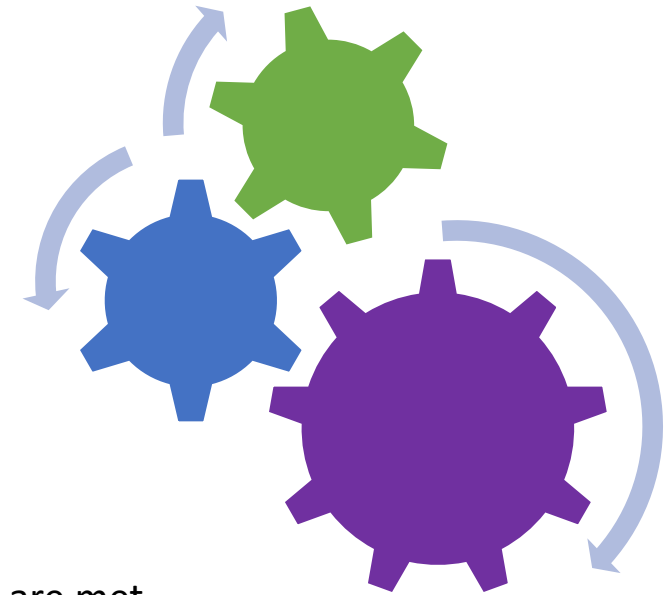
AI Ethics in Scholarly Communication

Publishers are involved with AI in many ways:

1/ Content, context & data providers

2/ Supporting internal workflows and services

3/ External-facing tools and services

STM
Advancing
trusted research

STM best practice principles for ethical and trustworthy AI

Transparency and Accountability
Community driven standards of transparency and accountability should be applied to the data used in AI input and training, as well as in the use of AI technology in publisher's tools, processes and services. Publishers encourage working with other actors to adapt standards where appropriate.

Sustainable development
AI systems are ideally positioned to address areas of global concern like climate change. Funding and other incentives for suppliers of high-quality input data can help maintain the vital supply of actionable knowledge.

Quality and Integrity
Quality and Integrity should be at the heart of the AI lifecycle, from the design and building of algorithms, to inputs used to train AI tools and services, to the application of AI. An appropriate IP framework is essential.

Fairness
To avoid bias, discrimination and the suppression of novel ideas, data selection and the application of AI must be carefully analyzed, planned, reviewed, and continuously monitored. Feedback mechanisms should be developed to report bias.

Privacy and Security
Principles that focus on data protection, data privacy and security can and should be used to respect and uphold privacy rights, data protection and ensure the security of datasets used in training or operating AI systems.

Legal and policy framework
The further development of AI should be guided by and grounded in clear legal standards and sound ethical principles. AI policy should respect intellectual property and continue to incentivize investment in high-quality content, datasets, and curated databases that can be used in AI applications. Any AI-enabling policy framework should foster the development of community-based standards and, where possible, build on existing initiatives. Due to the fast changing nature of the AI landscape, inflexible legislative tools should be avoided.

How STM helps combatting papermills

- Developing standards and best practices (in collaboration with COPE) in detection & handling

- Roll-out of principles and standards, and engaging with multiple stakeholders

- Create and grow awareness

- Hub to collaborate on technology

Thank you!

Rossum@stm-assoc.org

www.stm-assoc.org

STM Advancing trusted research