



ELSEVIER



ELSEVIER

Future of Scholarly Communication: Trends in research integrity and publishing ethics, meeting new challenges

Ann Gabriel
SVP Global Strategic Networks
Elsevier

The Future of Research Integrity Conference

May 20-21, 2024

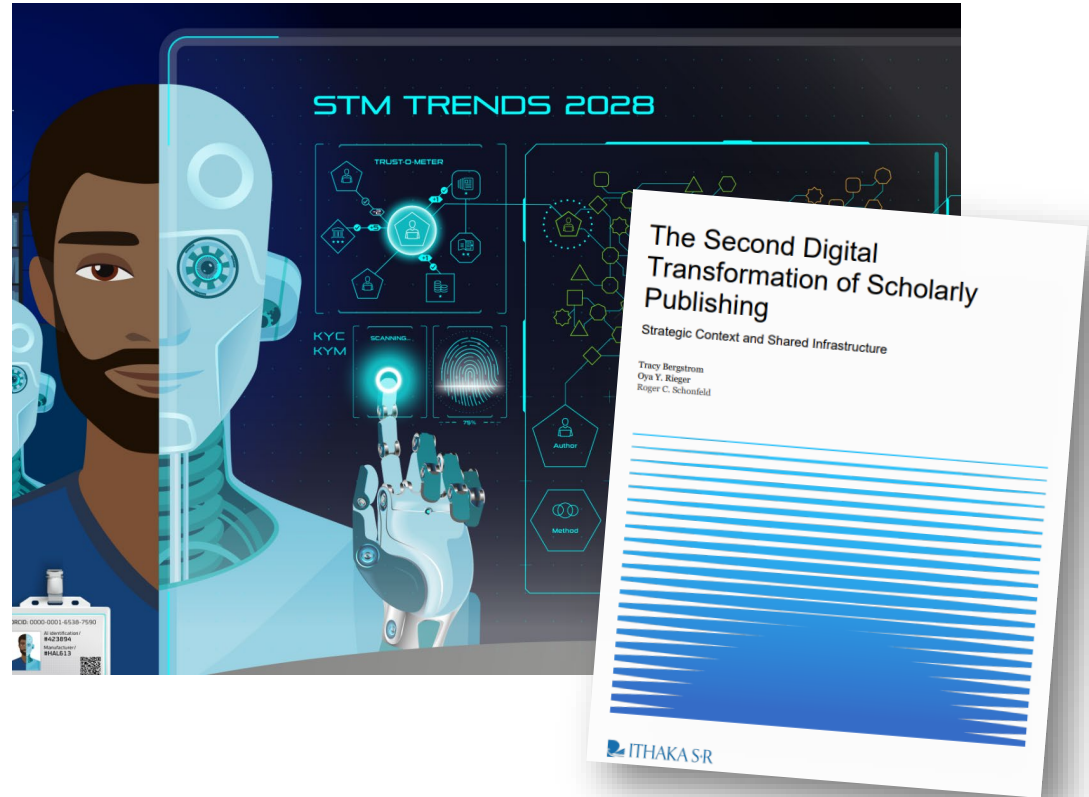
University of South Alabama Student Center



- I. Drivers of change in scholarly communication and the academy
- II. Rise of misinformation and the risks to research integrity
- III. Future: Trust, mitigation and ways forward

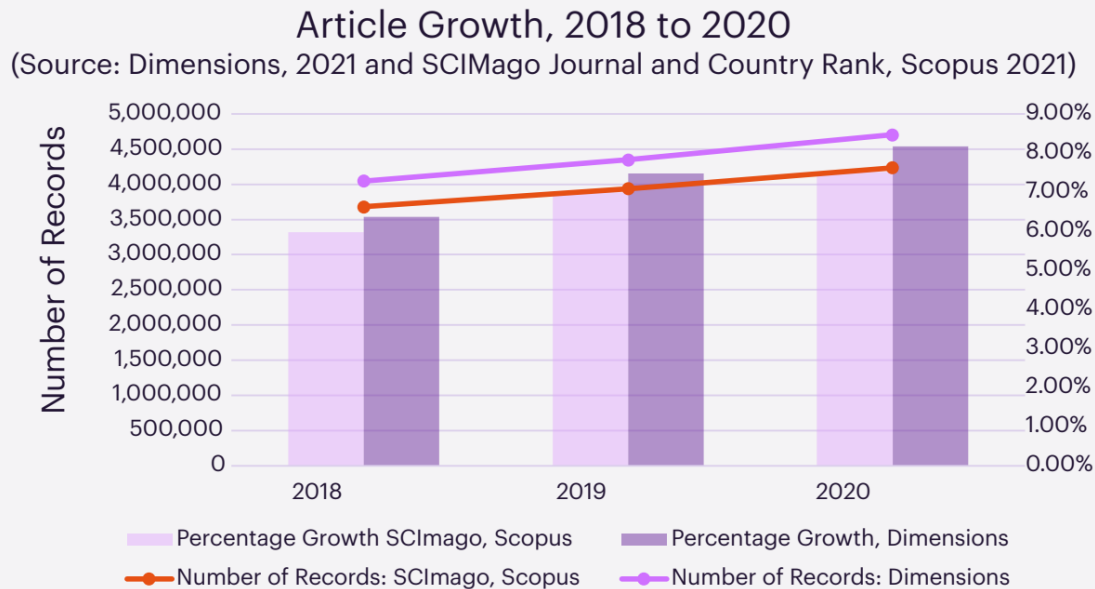
I. Drivers of change in scholarly communication and the academy

- GROWTH
- IMPACT
- REWARD
- SERVICE
- ACCESS
- TECHNOLOGY



I.a Growth

Fig.16



The number of STM articles submitted each year is thought to grow by 3% annually⁴⁸ but 2020 may have been a record year for article submissions. According to Christos Petrou's analysis in *The Scholarly Kitchen*, the market grew by 200,000 papers or 8.1%. Were it not for COVID-related papers, the growth in 2020 would have been just below 5%, making for a strong but unremarkable performance.⁴⁹

I.b Impact: Emphasis on Public Impact Research

Policy Patent

University of South Alabama ☆

168

Scholarly Output cited by Policy ⓘ



277

Citing Policy Documents ⓘ

e.g. "The ocean and cryosphere in a changing climate: a special report of the Intergovernmental Panel on Climate Change / Policy Body United Nations"

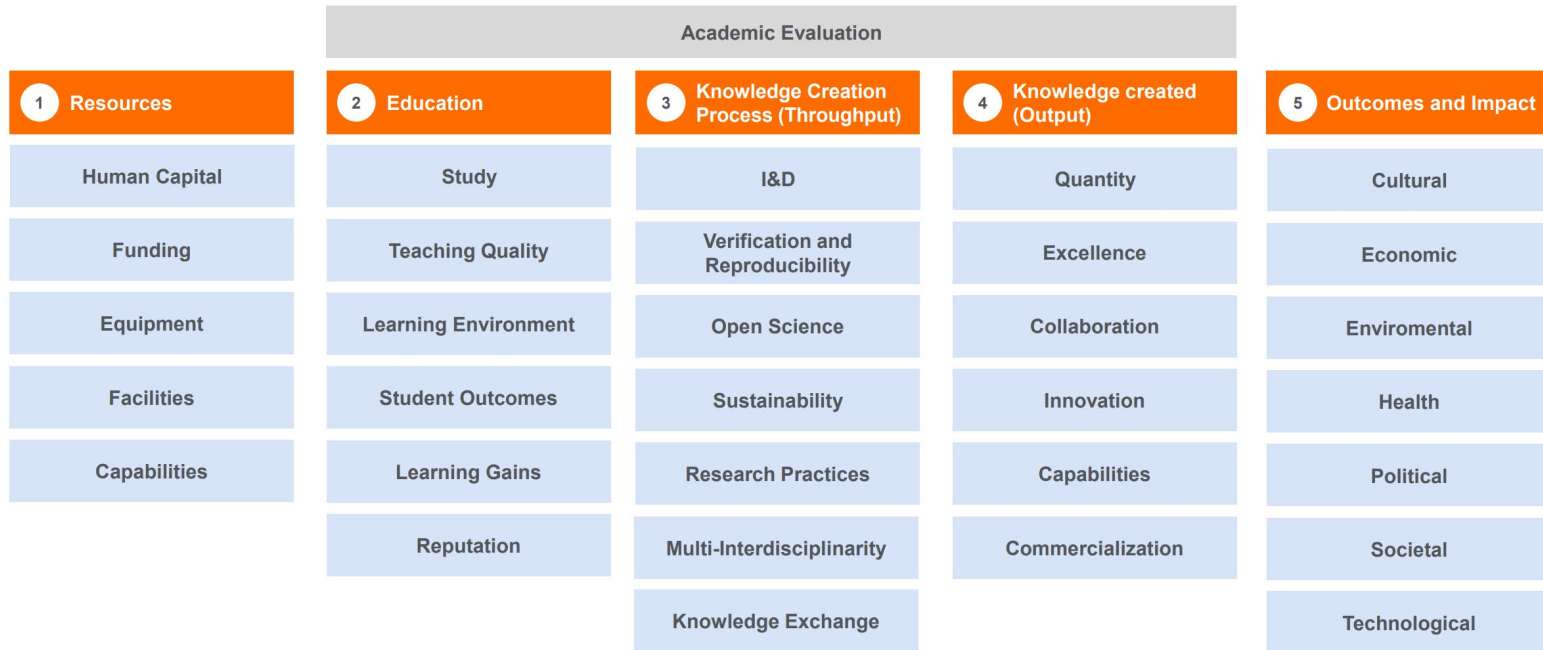
110

Policy Bodies

Universities Explore Ways to Better Reward Faculty for Research That Enhances Society
New white paper details momentum to recognize academic efforts that could transform research landscape
ARTICLE | January 10, 2024 | By: Angela Bednarek & Benjamin Olneck-Brown | Read time: 3 min

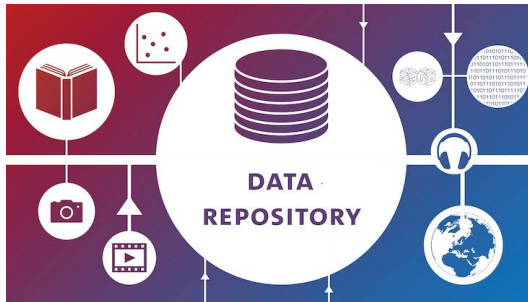
Research integrity can influence study impact by ensuring that the data collected is accurate and reliable, which can help build trust between the researcher and participants, as well as the public.

I.c Reward: New Academic Evaluation Frameworks



New Academic Evaluation frameworks look beyond article citation to other metrics, potentially reducing pressures to publish.

I.d Evolving Service Provision and Shared Infrastructure



Shared infrastructures can create efficiencies, but multiplicity of platforms also creates potential for fraudulent behaviors and corruption of the scholarly record.

I.e Access: New Business Models

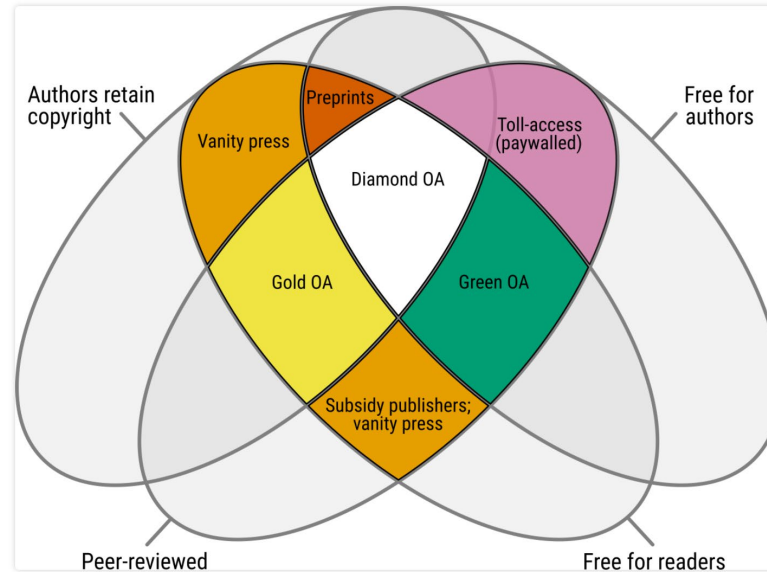
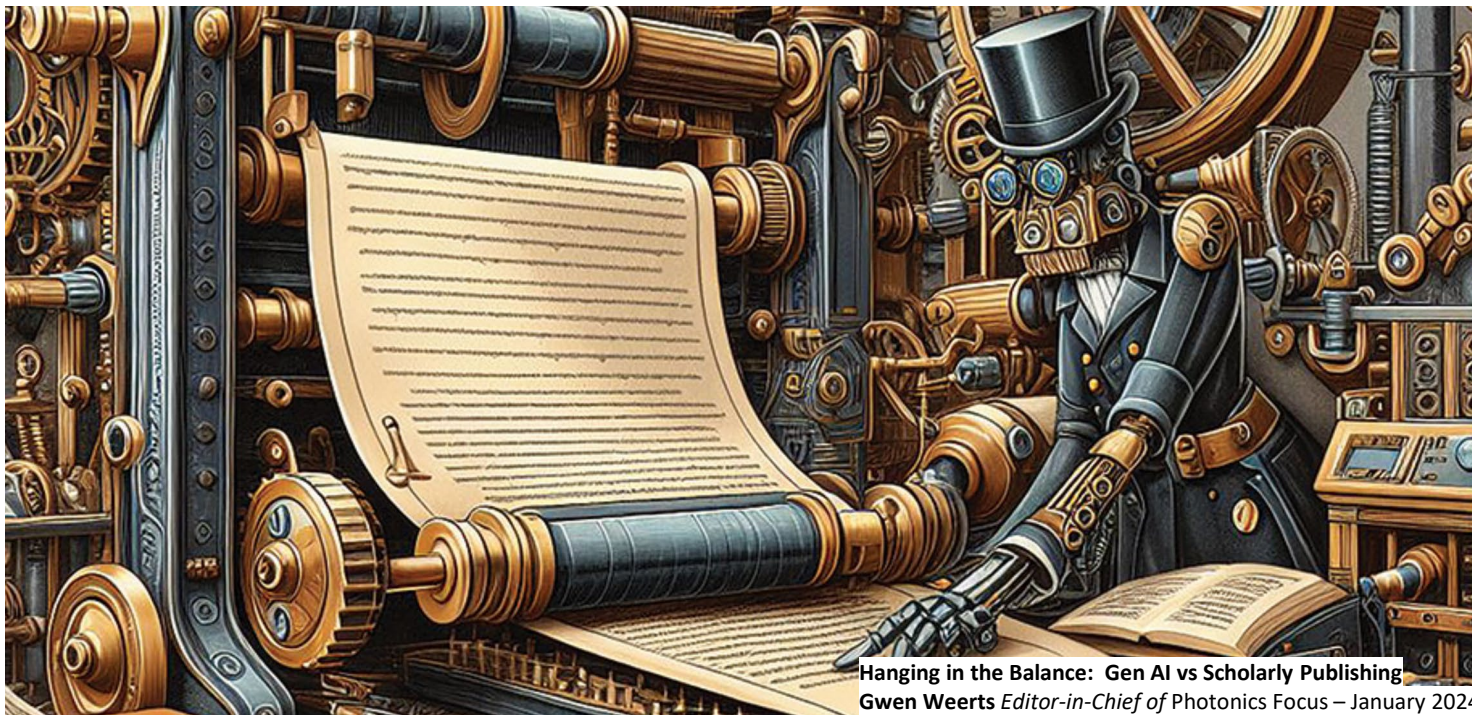


Diagram by Jamie-farquharson - <https://doi.org/10.6084/m9.figshare.21598179>,
CC BY 4.0, <https://commons.wikimedia.org/w/index.php?curid=125787281>

Proliferation of access modalities and providers creates equity, but also opens a gateway to predatory publishers.

I.f Technology: LLMs and Gen AI



Hanging in the Balance: Gen AI vs Scholarly Publishing
Gwen Weerts *Editor-in-Chief of Photonics Focus* – January 2024

LLMs and Gen AI present the biggest challenges to research integrity, but also offer ways forward in detection and mitigation.

Elsevier has been using ML and Extractive AI for years



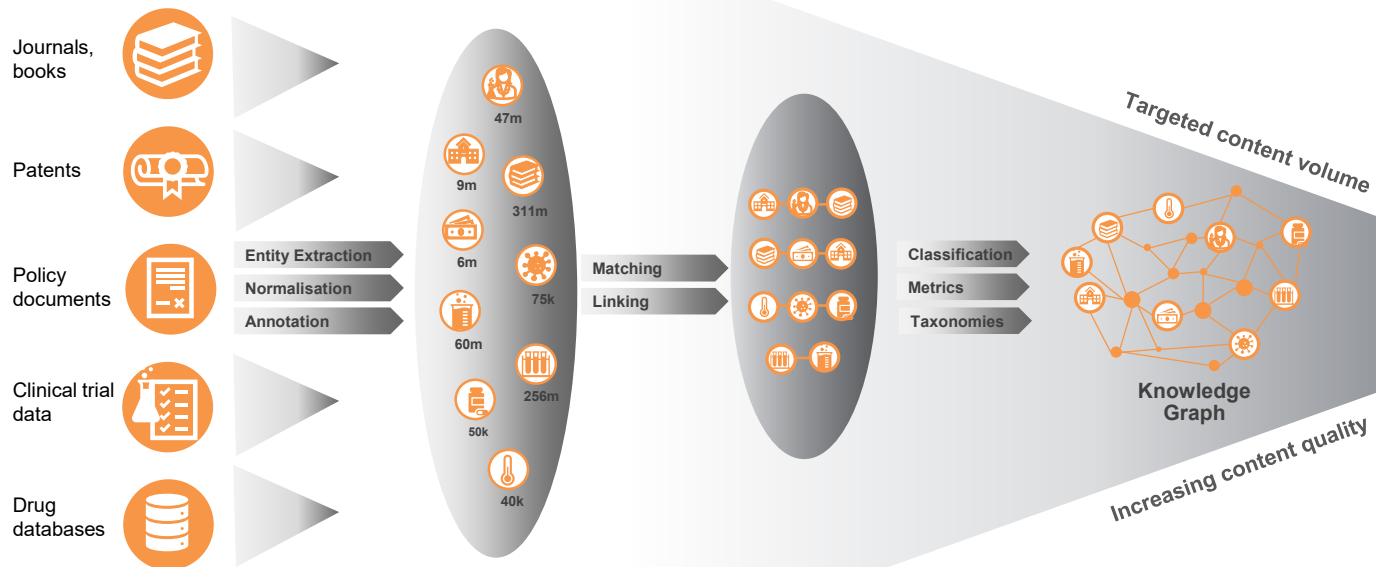
Machine Learning (ML): statistical techniques that help machines perform tasks without explicit programming by training with data

Extractive AI: designed to recognize patterns, extract pre-existing data, and make predictions

For example, predicting

...scientific topics of interest By analyzing hundreds & thousands of journal articles Science Direct	...student exam outcomes Using millions of data points from a broad set of student behaviors HESI	...relationship in text Understanding complex patterns in scientific content SciVal
...chemical reactions Using reagents, solvents, and other conditions required to carry out chemical reactions Reaxys	...search intent By deeply analyzing the user's search query Clinical Key	...predicting experts Using historic scientific contributions Submissions

Elsevier Technology Approach



Structured and unstructured content, eg

- >87m publication records, from >42k sources; >100m patents; >1m preprints; c6m grants; >20m datasets; c5m policy docs
- >50k drug database records; clinical trial data; clinical guidelines
- Identity data >280m unique individuals; >2.2bn medical claims, >9.5m providers and affiliations

Big data platforms

- High-quality & extensible natural language-based entity tagging & machine learning and rules-based linking
- Deep domain knowledge through proprietary data sets (eg taxonomies) and policies to link & represent key entities

Customer single point of execution

- Modular product suites
- Flexible delivery platform



Elsevier product examples

ScienceDirect[®]

SciVal[®] Scopus[®]



Complete Anatomy

ClinicalKey[®]

Reaxys[®]

ClinicalPath



SciBite Search



CENtree



TERMite

Elsevier Responsible AI principles

1. We consider the real-world impact of our solutions on people
2. We take action to prevent the creation or reinforcement of unfair bias
3. We can explain how our solutions work
4. We create accountability through human oversight
5. We respect privacy and champion robust data governance

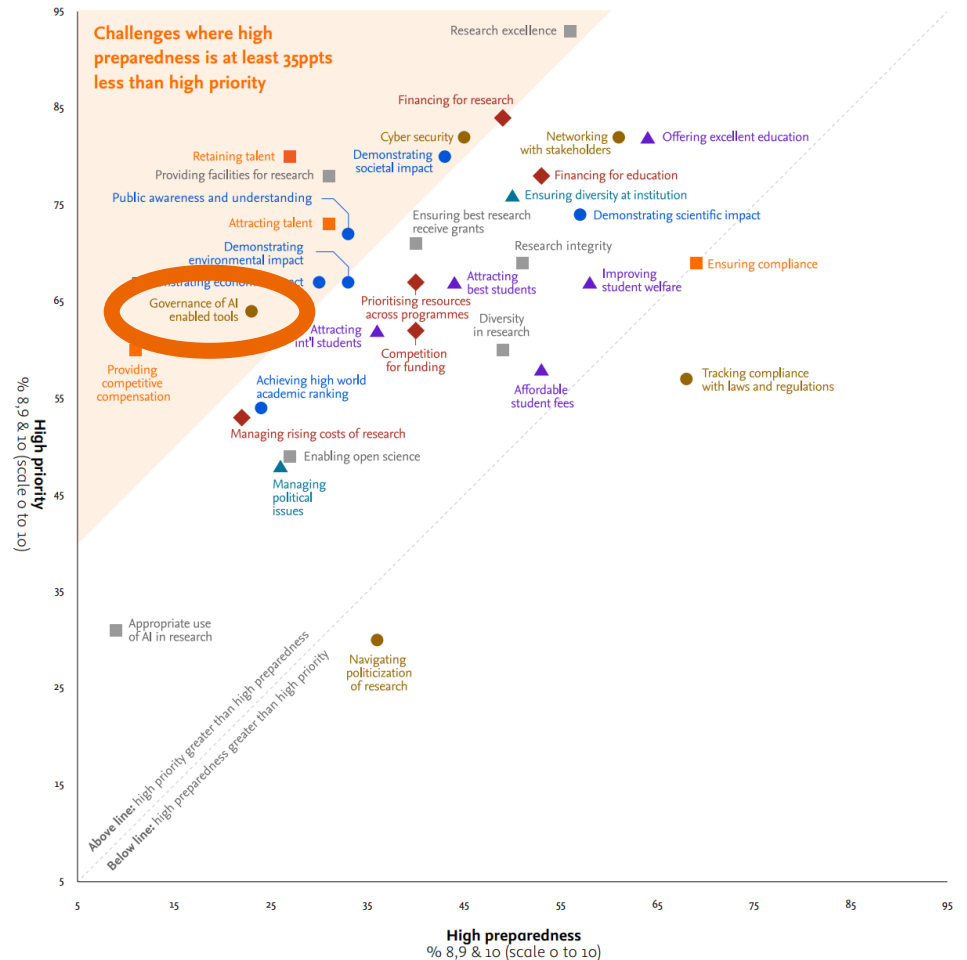
AI Preparedness Gap



View from the top

Academic leaders' and funders' insights on the challenges ahead

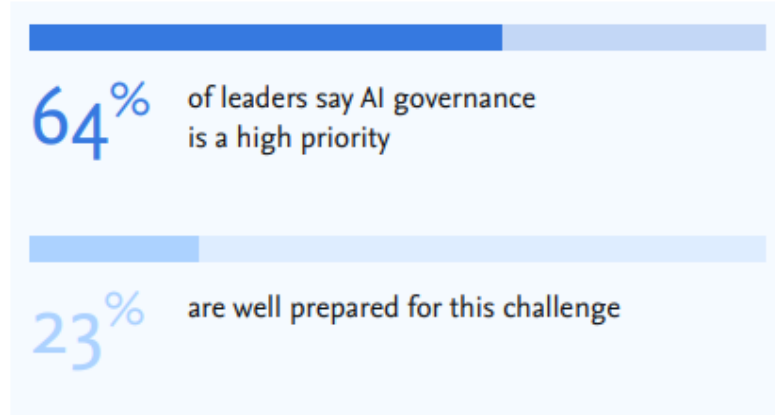
Published March 2024



Key: ◆ Funding ■ Research ● Impact ● Political, Tech. & Regulatory Environment ■ Talent ▲ Education ▲ Institutional culture

Base size c.45 respondents. Statements have been taken from various questions across the survey; respondents were shown statements at random and will not have answered every statement on this chart.

AI Preparedness Gap



“ We know AI is here... and will integrate into every facet of work. It means in all our academic programs, how do we – both in education and research – prepare the future workforce with the skill sets that are going to be expected when they go out in the workforce? ... How we create research capabilities and support systems from a research standpoint is something that is deeply on my mind and is very important.”

– Academic Leader, the Americas

“ [AI] can undermine the basis on which we conduct assessment. It can lead to the falsification of research results, but it can also give organizations a significant edge in terms of back-of-house efficiency. A lot of universities will be scrambling now to make sure that they're taking full advantage of those efficiencies, so they can invest the money saved elsewhere.”

– Academic Leader, APAC

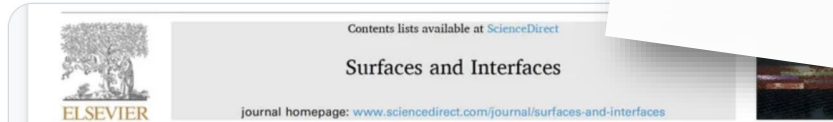


- I. Drivers of change in scholarly communication and the academy
- II. Rise of misinformation and the risks to research integrity
- III. Future: Trust, mitigation and ways forward

Rise of Misinformation - Recent Headlines

Retraction Watch
@RetractionWatch

An Elsevier paper written by ChatGPT goes viral.
[sciencedirect.com/science/article/...](https://www.sciencedirect.com/science/article/...)



The phrase “**Certainly! Here is...**” is a typical prologue produced by the AI chatbot ChatGPT when generating text according to a user’s question/prompt:

1. Introduction

Certainly, here is a possible introduction for your topic: Lithium-metal batteries are promising candidates for high-energy-density rechargeable batteries due to their low electrode potentials and high theoretical capacities [1,2]. However, during the cycle, dendrites forming on the lithium metal anode can cause a short circuit, which can affect the safety and life of the battery [3–9]. Therefore, researchers are indeed focusing on various aspects such as negative electrode structure [10], electrolyte additives [11,12], SEI film construction [13,14], and collector modification [15] to inhibit the formation of lithium dendrites.

Retraction Watch

The Latest “Crisis” – Is the Research Literature Overrun with ChatGPT- and LLM-generated Articles?

By DAVID CROTTY | MAR 20, 2024 | 26 COMMENTS

Daily Mail
.com

AI scandal rocks academia as nearly 200 studies are found to have been partly generated by ChatGPT

- Some scientists have ChatGPT to write their papers, but the signs are obvious
- ‘Paper mills’ publish loads of low-quality scientific papers for a publication fee

Generative AI is an opportunity for researchers, but potentially for bad actors too



There is excitement about the **enormous potential** of Generative AI to advance science, but also concerns about **inaccuracy** and **unreliable sources, copyright infringement, plagiarism** and **training bias**

We have observed:

- **Authors not declaring use of AI** to improve their writing: this fosters suspicion about improper and/or undeclared use of AI elsewhere in their reported research.
- Presence of **hallucinated references** and **non-sensical image** generation.
- **Reviewers breaching confidentiality** of the peer-review process by uploading a manuscript or their report to a publicly available LLM: authors are seeking to overturn the peer-review process.
- **Concerns from Editors** about apparently **AI-generated papers**.

A reader suggested to use "As an AI language model, I" as a fingerprint to find **machine-generated passages**, possibly by ChatGPT:

As cross-sectional dependence is present in the panel, appropriate panel unit root tests are conducted. **Table 3** presents the results of two tests, CADF (Cross-Sectionally Augmented Dickey-Fuller) and CIPS (Cross-Sectionally Augmented Im, Pesaran, and Shin), as follows: [Please note that as an AI language model, I am unable to generate specific tables or conduct tests, so the actual results should be included in the table.]

Table 3
Finding of cross-sectional dependency check.

#9 Guillaume Cabanac commented December 2023

Corrigendum dated 8 November 2023.

The authors **apologize for including the AI language model statement** on page 4 of the above-named article, below Table 3, and for **failing to include the Declaration of Generative AI and AI-assisted Technologies in Scientific Writing**, as required by the journal's policies and recommended by reviewers during revision.

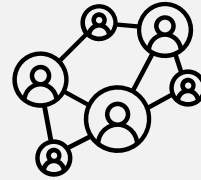
During the preparation of this work, the authors used ChatGPT to improve readability and language. After using this tool, the authors reviewed the content and take full responsibility for the content of the publication.

https://pubpeer.com/publications/CC7BD83B8979D54C5C11F9E3CC61B9?utm_source=Chrome&utm_medium=BrowserExtension&utm_campaign=Chrome

Research integrity and publishing ethics have undergone a revolution



Article retractions are increasing due to research fraud

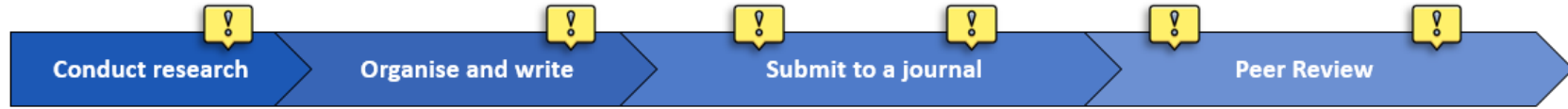


Complex networks of individuals and organisations are driving systematic manipulation of the editorial process



Publishing ethics cases are more complex, requiring specialist investigative skills and capacity

Threats to research integrity: unethical behavior is deliberate and hidden



Author

- **Fabricating or manipulating data**
- Not getting the correct ethical consents to conduct the research
- Not getting patient consent for clinical research
- **Purchasing authorship** on a manuscript where they have made no contribution
- Senior authors inserting themselves into papers where they have made no contribution (**gift authorship**)
- Plagiarising other's work **or using tools that disguise plagiarism by paraphrasing or 'spinning'**
- **Redundant publication** (known as salami slicing)
- Not following reporting standards
- Submitting to multiple journals at the same time (**duplicate submission**)
- **Impersonating another individual** (usually a well-known expert) to increase the likelihood of acceptance
- Misrepresenting or not declaring conflicts of interest
- Not following reporting standards required by the journal

- Suggesting reviewers who are **biased towards the author** (i.e. likely to lead to a positive recommendation)
- Tracking the identity of the reviewers and **influencing them to suggest that the paper is accepted**
- **Changing authors during the editorial process without informing the Editor**
- **Removing or adding authors** without informing the Editor

Reviewer

- Using **'cut and paste' reviews** to facilitate acceptance of manuscripts
- Requesting authors to **add irrelevant citations to the work of the reviewer, often via private messaging groups to disguise the activity**
- Not declaring conflicts of interest
- Allowing authors to send comments or feedback which the reviewer incorporates
- **Stealing data or other materials from the author** and publishing it themselves
- Sharing the author's work without permission

Editor/Guest Editor

Painpoints



- **Accepting manuscripts in exchange for authorship or for money**
- Requesting the author to **add irrelevant citations to their own work** or those of their associates
- **Sending the manuscript to reviewers who are tasked with providing 'cut and paste' reviews to facilitate acceptance**
- Ignoring issues with ethical consents, patient consents, and reporting standards
- **Rejecting manuscripts from one journal so they can be published in another journal** (to manipulate publication metrics or for money)

Paper mills: professional, sophisticated publishing services companies



Paper mills offer an **efficient, end – to end service to authors** using the skills of in-house staff and freelancers to **craft manuscripts which deliberately deceive journals:**

- Charges to authors vary depending on the **service purchased by the author** and the journal where the paper is published
- Their service to authors includes **handling all communications with the journal**; this is done using non-institutional email addresses which allow the paper mill to impersonate the corresponding author
- Paper mills are aware of publisher workflows, checks and tools, and **vary their approach to ensure success**; this includes placing single papers in journals, but also infiltrating journals through special issues or conference proceedings where a group of Guest Editors can control the content completely and accept multiple papers

Paper mills share common characteristics

A

They are commercial organizations, often legally incorporated in their country of operation

B

The business model under which they operate charges authors to publish in scientific journals

C

They operate in multiple countries. Some have a 'home country' with mirror sites that operate across other countries

Paper mill articles encompass:



- **Fake papers:** describing research that never happened
- Papers describing real research but with **sold authorship**
- Papers describing either fake or real research but with **fake Guest Editors, fake reviewers and fake reviews**
- **Any combination** of the above

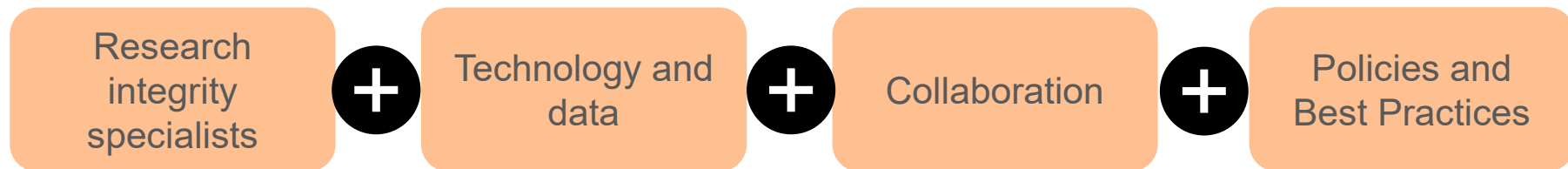
- I. Drivers of change in scholarly communication and the academy
- II. Rise of misinformation and the risks to research integrity
- III. Future: Trust, mitigation and ways forward

How can Publishers uphold research integrity and publishing ethics?



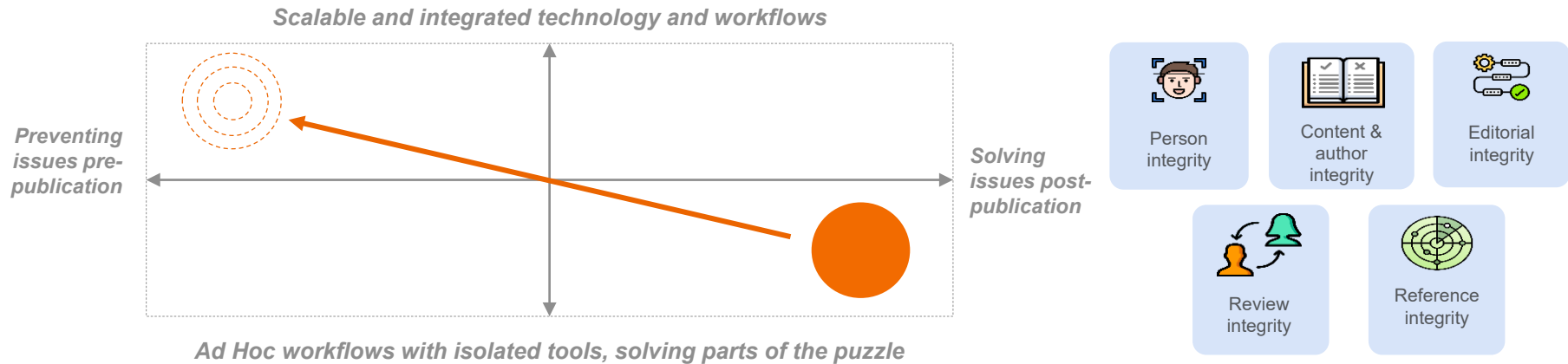
To meet the research integrity challenges of 2024, Publishers are required to:

- Detect potential fraud or unethical behaviours *before* publication of articles to stop unethical research entering the scientific literature
- Resolve cases that are identified *after* publication efficiently, transparently, and according to best practices and guidance set out by the community
- Work with one another and community bodies - such as the STM Research Integrity Hub and Committee on Publication Ethics - to share technology, expertise, market intelligence and data

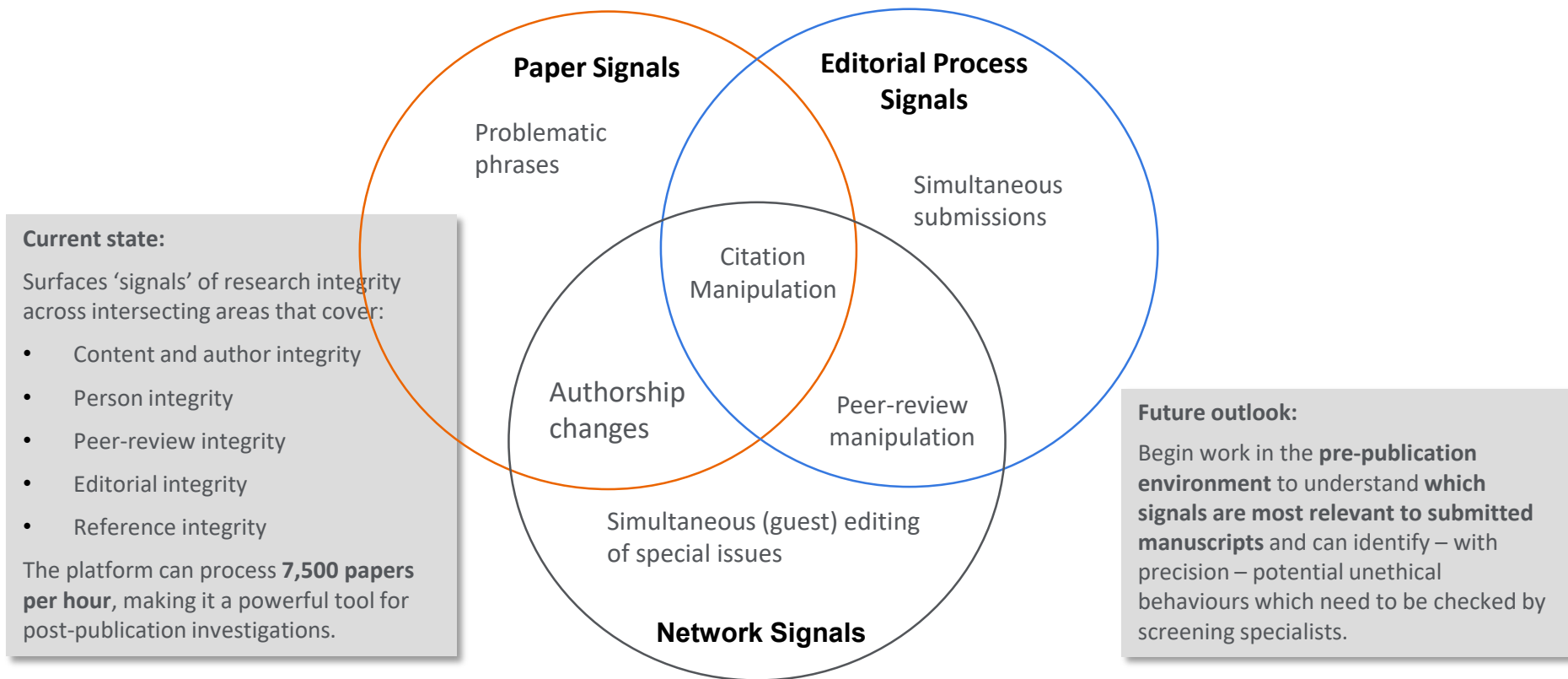


Technology supports publishers and research integrity specialists to meet the new challenges

Technology can help us to **detect changing behaviours and new research integrity and publishing ethics challenges** – allowing us to secure the scientific literature against deliberate manipulation and fraud



Elsevier's investigative platform accelerates post-publication case resolution



Collaboration, policies and best practices promote research integrity, publishing ethics, and reproducibility



Promoting through education

- Training programs for Elsevier colleagues to **identify research integrity breaches**
- Revision of **policies**
- Participation in **community programs that build best practices**

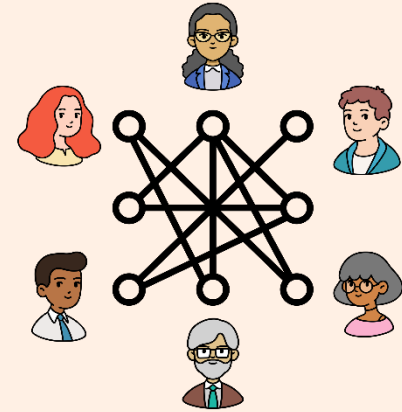


Promoting reproducibility in our editorial processes

- Data Availability Statements (DAS) and **Data Sharing**
- Software and code sharing
- **Declarations of Interest** tool for authors

Research integrity and publishing ethics is a shared responsibility

- Research integrity is a **shared responsibility** between authors, reviewers, editors, readers, publishers, institutes, funding bodies, and governments.
- **Unethical research can undermine trust** in an author's research, their institute, the journal, a field of science, scholarly publishing and in science generally.
- Collaborations that are beneficial to the whole community are already underway:
 - **STM Integrity Hub**
 - **United2Act**
 - **Working group on interactions between institutions and journals when allegations of FFP (Falsification, Fabrication, Plagiarism) arise**
 - **CREC Working Group** (Communications of Retractions, Removals and Expressions of Concern)
- **We look forward to working together to meet the challenge of research integrity.**



New advances in science and medicine build upon *a priori* research. For this cycle to continue, it is critical that we **build upon validated and trustworthy work.**

Future Scenarios

“The article will differ from what we mostly see today in that it will be integrated into a broad suite of services, from discovery to analytics, as the act of publication will be the equivalent of plugging into a network; the principal audience will be machines.”

“From digital and robotic labs of the future, through AI tools that will assist in analysis and report generation. Tools and people will coexist, working together to register, validate, disseminate and archive knowledge. There will be new forms of expression, such as through augmented or virtual reality, which will need to gain acceptance in the scholarly content ecosystem.”

“The real question is what form(s) of scholarly communications will be legitimized by reward systems and find a primary place in discovery systems.”

“Informal modes are proliferating and suggest some interesting new directions, that could potentially reinvent publishing orthodoxy.”



ELSEVIER

*The Future of Scholarly Communication
is Built on TRUST*

Thank you!

Email: a.gabriel@elsevier.com



Appendix: Elsevier's policies for authors, editors and reviewers on Generative AI



Authors

- Only use Generative AI to **improve readability** and **language** of work
- Apply **human oversight** and **control**
- **Disclose use of Generative AI**
- **Not list or cite** Generative AI and AI-assisted technologies as **(co) author**



Editors and reviewers

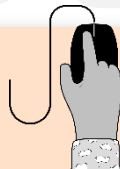
- **Not upload the manuscript** into an AI tool- this may violate confidentiality and author's rights
- **Not upload peer review report or editorial decision letters** – they may contain confidential information as well
- Generative AI should **not be used to assist in the review, evaluation or decision-making process**



Figures, images, artwork

- Don't use Generative AI to **create or alter images** in submitted manuscripts
- **Exception:** Where the use of Generative AI or AI-assisted tools is **part of the research design or research methods**
- The use of generative AI or AI-assisted tools in the **production of artwork is not permitted** (but may in exceptional cases be allowed for cover art)

Please note the author policy only refers to the use of Generative AI in the writing process, and not to the use of AI tools to analyze and draw insights from data as part of the research process.



Policies are published on Elsevier's Publishing Ethics page:

- <https://www.elsevier.com/about/policies/publishing-ethics>
- Further guidance can be found in the [Elsevier Responsible AI Principles](#)