

# Research integrity and open access models: insights from Retraction Watch and OpenAlex

This study examines the relationship between research integrity and open access (OA) publishing models using data from Retraction Watch and OpenAlex. Analysing 60,608 retracted publications from 2009 to 2024, the article traces how retraction patterns have shifted alongside the expansion of OA, with gold OA surpassing closed access as the dominant modality among retracted articles by 2023. The analysis also highlights the economic dimensions of research integrity and OA, with an estimated US\$41.9 million in article processing charges (APCs) collected by publishers for research that was later retracted. These findings raise concerns about APC-based publishing models that directly link publisher revenue to publication volume, creating structural tensions for editorial oversight and quality control. Rather than framing OA as inherently more or less prone to integrity failures, the article argues that these challenges reflect broader incentive structures within contemporary scholarly publishing. Addressing them will require co-ordinated governance efforts among publishers, funders, libraries and research institutions to ensure that OA is matched by accountability, transparency and trust in scholarly research.

## Keywords

research integrity, open access, article processing charges, scholarly publishing

## Introduction

Over the past several decades, open access (OA) has moved from the fringes of the scholarly publishing ecosystem to the forefront. Advocates of OA highlight its role in democratizing knowledge, accelerating innovation and increasing the access to and impact of research. Increasingly, funders, higher education institutions and government agencies are developing and implementing OA policies and mandates for publicly funded research. Publishers, on the other hand, have responded with a wide variety of OA business models that include Read & Publish agreements (also known as transformative agreements), gold OA, hybrid OA and Subscribe to Open (S2O). Alongside these shifts in the scholarly publishing landscape, questions of research integrity – defined by the UK Research Integrity Office as ‘encompass[ing] all aspects that support ethical and responsible research practices, which help to foster trust and confidence in the research process. It applies to all disciplines and sectors where research is conducted and includes everything from the initial concept and design of a project to its execution and the dissemination of results and findings’ – have intensified ([UK Research Integrity Office, 2026](#)).

Research integrity issues such as data fabrication and data manipulation, and other forms of misconduct, are being tracked more systematically now than ever before and have highlighted some of the vulnerabilities in the peer-review process as well as the incentive structures that shape scholarly publishing, particularly as they relate to promotion and tenure. While retractions represent a small fraction of the scholarly literature, they play a critical role in maintaining research integrity by correcting and signaling where breakdowns in research or editorial processes have occurred. At the same time, retractions can have significant implications for trust in scholarly communication, highlighting the importance of strong practices in transparency, accountability and oversight.



BEN RAWLINS

Associate Dean  
for Outreach,  
Engagement and  
Collections  
University of  
Kentucky  
United States

‘While retractions represent a small fraction of the scholarly literature, they play a critical role in maintaining research integrity’

While there are many potential factors that contribute to the rising concern of research integrity – including generative artificial intelligence (genAI), predatory publishers and paper mills, among others – the connection and intersection of OA and research integrity is an area that requires further exploration. A benefit of OA is that it increases the visibility of research, which could support the faster identification of errors, misconduct or other problematic research practices. Conversely, the growth of and reliance on article processing charge (APC)-based publishing models raises concerns around financial barriers for underfunded researchers, predatory journals and incentivizing publishers to prioritize quantity over quality. Understanding how research integrity issues manifest across different OA modalities is essential not only to increasing access to research but also to ensuring that research adheres to the highest standards of research integrity. Two openly available datasets that enable this exploration are Retraction Watch and OpenAlex. Together, these resources allow for an in-depth examination of how retractions intersect with different OA modalities across disciplines and over time. Using these resources, this study seeks to examine questions such as the following:

'Understanding how research integrity issues manifest across different OA modalities is essential'

- What patterns emerge when retracted publications are grouped by OA modality?
- How do these patterns and trends vary across disciplines and publishers?
- What insights can libraries, funders, higher education institutions, publishers and policymakers draw from this analysis to ensure both research integrity and sustainable OA practices?

By addressing these questions, this study aims to advance conversations that can continue to push for the creation of a scholarly publishing ecosystem that both increases access to research and remains reliable and trustworthy.

### Literature review

While the expansion of OA publishing has made research more widely available, particularly through gold and hybrid models, it has also raised new concerns about research integrity, editorial quality and how journals handle corrections and retractions. These concerns are especially pronounced in contexts where article processing charges (APCs) – a hallmark of gold and hybrid models – shape publication decisions or where increased visibility complicates post-publication oversight ([Björk & Solomon, 2012](#); [Zheng & Fu, 2024](#); [Chen et al., 2021](#)). This shows that while OA has expanded access to research, it has also introduced new integrity challenges that are deeply entangled with the economic structures at the heart of gold and hybrid OA publishing.

The pricing structures of APCs in gold and hybrid OA reflect a complex power dynamic where commercial publishers are able to leverage the high value that researchers and institutions place on journal prestige, as well as compliance mandates from funders, to drive and maintain revenue. Hybrid journals, which operate using a dual revenue model of subscription fees and APCs, consistently charge significantly higher median APCs than gold OA journals, suggesting that publishers strategically set these fees to maximize revenue through 'double dipping' (something they deny), thereby generating income from both subscriptions and author payments ([Tocco et al., 2025](#); [Solomon & Björk, 2016](#)). This system provides commercial publishers with immense power and authority to dictate pricing, which has led to APC hyperinflation at rates that far exceed the rates of inflation. This is because authors, particularly those from well-resourced institutions and with adequate grant funding, tend to be price-insensitive and prioritize journal prestige over OA publishing costs ([Khoo, 2019](#)). This shapes who is able to participate and magnifies already prevalent inequities, as researchers from well-resourced institutions or with sufficient grant funding are significantly more likely to choose paid OA venues, while researchers from under-resourced institutions or lacking sufficient funding – particularly those from low- and middle-income countries – face prohibitive financial barriers that steer them toward lower-quality publishing outlets or subscription

'The pricing structures of APCs in gold and hybrid OA reflect a complex power dynamic'

3 journals where waivers are often limited or, in some cases, unavailable (Vervoort et al., 2021; Ellingson et al., 2021; Tocco et al., 2025; Siler et al., 2018). While there have been various attempts by funders, institutions and libraries to constrain costs through transformative agreements (TAs) and APC price caps, the current market conditions enable publishers to continue to set high APCs, which effectively ensures that only those with adequate financial resources can actively participate without cost being a barrier (Jahn & Tullney, 2016; Ellingson et al., 2021; Khoo, 2019). As a result, APC-based OA models do more than influence publishing costs; they actively shape who can participate in scholarly communication and under what conditions, raising critical questions about equity, access and governance in the OA publishing ecosystem.

'APC-based OA models do more than influence publishing costs; they actively shape who can participate in scholarly communication'

While the economics of OA, particularly APC-based models, raises concerns about financial barriers and incentives, research also suggests that OA status can have positive effects after publication. Research outputs that are OA help to increase transparency and broaden dissemination, which in turn can help surface errors or misconduct quicker (Zheng & Fu, 2024). Empirical analyses support this pattern, showing that gold OA publications have both the highest retraction rates and the shortest average time to retraction (2.95 years) (Zheng & Fu, 2024). This finding is open to at least two interpretations that are not mutually exclusive. On the one hand, it may suggest that gold OA adds value to research integrity by enabling errors to be identified and corrected quicker than in green OA or non-OA publications, likely because broader accessibility exposes articles to greater post-publication scrutiny. On the other hand, the elevated retraction rate for gold OA may also reflect the consequences of lax acceptance standards in APC-funded journals, where the conflict of interest between publisher revenue and editorial rigor is most acute. Distinguishing between these explanations is difficult without controlling for article volume across publishing modalities. These benefits are only realized when post-publication governance is well-defined, and metadata are interoperable across publishers, repositories and indexing services (Bordignon, 2025). However, in many repositories, OA primarily supports dissemination rather than correction, resulting in potential blind spots where updates about retractions or corrections are missing or inconsistently applied. As a result, for integrity mechanisms to work as intended, there need to be clear, interoperable workflows and bidirectional metadata that signal changes in a publication's status and reliably link publisher notices to the version of record across repositories and platforms (Bordignon, 2025). These findings suggest that the contribution of OA to research integrity depends less on access alone and more on the strength of the post-publication infrastructure that ensures corrections and retractions are visible, consistent and trusted.

'OA publications have both the highest retraction rates and the shortest average time to retraction'

Maintaining research integrity in APC-based publishing models, according to the literature, requires a multipronged governance approach, especially where publisher income increases with article volume, creating potential conflicts of interest that can affect editorial decision-making (Solomon & Björk, 2016; Khoo, 2019). This conflict is not limited to legacy commercial publishers that have added OA options to existing subscription portfolios. Large OA publishers operating entirely on APC revenue at scale, most prominently MDPI, face structurally identical incentives: every rejected manuscript is lost revenue, while every accepted one generates income. At the volumes at which these publishers operate, even modest shifts in editorial thresholds can have significant consequences for the integrity of scholarly publishing. MDPI has faced sustained scrutiny over editorial quality, including concerns about peer-review practices, rapid acceptance timelines and the proliferation of special issues as a mechanism for volume expansion (Brainard, 2026). One important measure identified includes mandating price transparency through the public reporting of APC expenditures, as seen in the Open APC initiative, which facilitates the monitoring of pricing patterns and institutional spending (Jahn & Tullney, 2016). To curb APC

'Maintaining research integrity in APC-based publishing models, according to the literature, requires a multipronged governance approach'

4 hyperinflation and ensure affordability, some funders have implemented strict price caps, such as the €2,000 limit enforced by the Deutsche Forschungsgemeinschaft (DFG) or the caps explored by cOAlition S under Plan S ([Jahn & Tullney, 2016](#); [Ellingson et al., 2021](#); [Khoo, 2019](#)). However, price caps have their fair share of criticism, as noted in the public comments regarding the U.S. National Institutes of Health (NIH)'s proposed APC caps. Commenters suggested that the proposed caps do little to address deeper structural problems in scientific publishing and may disproportionately exclude some researchers from publishing in prestigious journals, which can in turn affect their chances of securing jobs, grants or promotion ([Jacobs, 2025](#)). Furthermore, fostering global equity requires transparent and, ideally, automated waiver policies to remove financial barriers for authors from low-resource settings or those lacking grant funding ([Vervoort et al., 2021](#); [Gasparyan et al., 2013](#); [Ellingson et al., 2021](#); [Siler et al., 2018](#)). Overall, the literature suggests that protecting research integrity and equity in APC-based publishing depends on co-ordinated and collaborative governance efforts among publishers, funders, libraries and platform providers that includes addressing pricing power, transparency and inclusion throughout the research life cycle rather than relying on one-off policy fixes.

## Methodology

The data for this study came from two sources: Retraction Watch and OpenAlex. Retraction Watch is a project of the Center for Scientific Integrity that tracks retractions of scientific papers to promote transparency and integrity in research. Their dataset, the Retraction Watch Database, is a comprehensive record of retracted publications, containing metadata on retraction reasons, journal and publisher details and links to original articles. OpenAlex is an open index of scholarly research developed by OurResearch that provides metadata on publications, authors, institutions, research topics and connections between these works. OpenAlex also includes metadata on the open access (OA) status (diamond, gold, green, hybrid, bronze or closed) of a work and, when available, data on article processing charges (APCs). OpenAlex was chosen for this analysis because it is an open alternative to proprietary databases like Scopus or Web of Science and has broad coverage across disciplines, detailed OA indicators and a flexible API (application programming interface).

The Retraction Watch dataset, available for download on GitHub ([Retraction Watch, 2025](#)), was accessed and downloaded on February 24, 2025, and again on May 8, 2025, to incorporate any additional retractions reported from 2024. Although the dataset includes retractions from 2025, these were excluded from the analysis, which focused on retractions from 2009 – the first year that retractions exceeded 1,000 – to 2024.

In the Retraction Watch dataset, there is a subject field that has a subject prefix followed by a more specific subject – for example, '(PHY) Chemistry'. In the modified dataset, two different Excel VBA functions were created (ExtractOutsideParentheses and MapCodesToDescriptions) and used to create an overarching subject domain (e.g. Physical Sciences) in one column and a subject subfield (e.g. Chemistry) in another column. The ExtractOutsideParentheses function scans a text string and extracts all content outside of parentheses, preserving multiple segments and separating them with semicolons. The MapCodesToDescriptions function takes semicolon-separated codes (e.g. SOC; PSY) and looks up each code in a specified sheet, returning the corresponding full-subject prefixes defined by Retraction Watch while removing duplicates ([Retraction Watch, n.d.](#)).

Using the OpenAlex API, Python scripts were utilized to obtain data by digital object identifier (DOI) or PubMed ID. These scripts pulled data from the following OpenAlex fields: oa\_status, publication type type\_crossref (this replaced the publication type in the Retraction Watch dataset) and apc\_paid (if available). For records that did not have

'protecting research integrity and equity in APC-based publishing depends on co-ordinated and collaborative governance efforts'

'the Retraction Watch Database, is a comprehensive record of retracted publications'

5 either a DOI or PubMed ID, the data from the `oa_status`, `type_crossref` and `apc_paid` fields were retrieved manually. Once the data were retrieved from OpenAlex, they were merged with the Retraction Watch dataset. Overall, the final dataset included a total of 60,608 publications with 22,827 OA publications identified as either gold or hybrid. To store and manage the data, they were populated into MySQL – an open-source relational database. Then, a website was created to analyse and visualize the data using other open tools and frameworks such as PHP, Bootstrap and HighCharts.

### Limitations

It is difficult to gain an accurate picture of the APCs paid because APC data are often not readily available due to the lack of transparency in the market. The APC data in OpenAlex come from OpenAPC, Directory of Open Access Journals (DOAJ) and publisher websites but do not include data on all gold and hybrid OA publications. For example, APC data were not available for 1,859 OA publications; therefore, the data on APCs paid for retracted research are a rough estimate and are likely an underestimation.

### Findings

Between 2009 and 2024, there were a total of 60,608 retractions. While retractions were already present at non-trivial levels at the beginning of the period, numbering just over 1,000 in 2009, the total number of retractions expanded over time, with periods of relative consistency punctuated by sharp increases, particularly in more recent years. By the early 2020s, annual retraction counts had reached levels several times higher than those in the late 2000s and early 2010s, underscoring how retractions have become a far more prominent, and consequential, occurrence within scholarly publishing.

Figure 1 shows the trends in the number of retractions from 2009 to 2024. Following an initial spike around 2010–2011, retraction counts declined and remained comparatively low and consistent throughout the mid-2010s. However, beginning around 2017, the trend began to shift upward, with steady year-over-year increases in retractions through 2022. This upward trajectory reached its peak in 2023, when retractions more than doubled compared to just two years earlier, before declining in 2024. Although the final year shows a reduction from the 2023 high, retraction levels remain well above those recorded for most of the preceding decade. Taken together, the data in the figure highlight both the long-term increase in retractions and the increasing volatility of retraction activity in the most recent years.

'retractions have become a far more prominent, and consequential, occurrence within scholarly publishing'

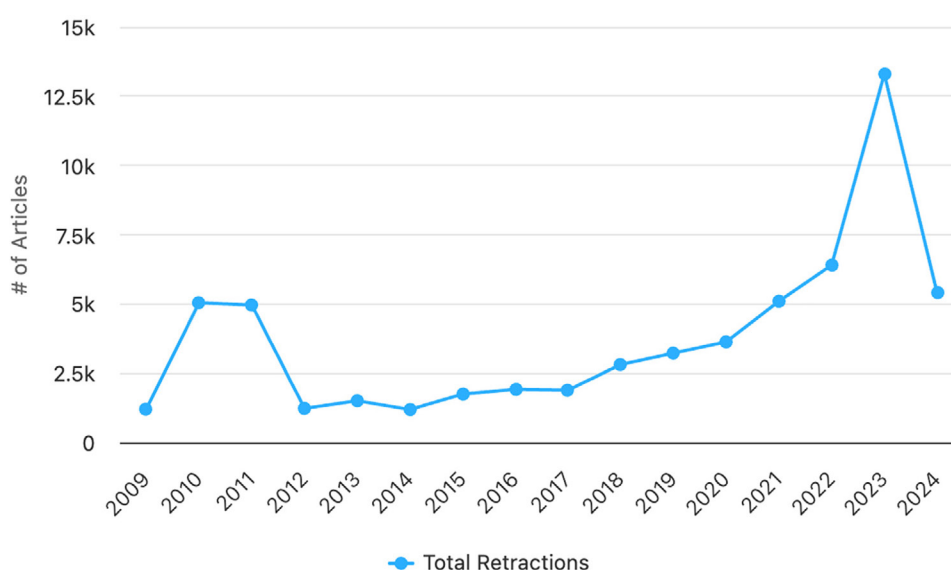


Figure 1. Retractions by year, 2009–2024

## Retractions by access type

Scholarly research is published and disseminated under a range of different access options, including traditional subscription-based models that are paywalled or OA models such as gold and hybrid. OpenAlex enables the classification of scholarly works – in this case, retracted research – into several access categories that reflect how and where those research outputs are made available. These include closed access works that remain behind paywalls in traditional subscription models; gold OA, where articles are published in fully OA journals; hybrid OA, where individual articles are made open within subscription journals, typically through APCs; green OA, where versions of works are made available through repositories; bronze OA, where articles are free to read on publisher platforms but lack a clearly defined reuse license; and diamond OA, where articles are published in a fully open journal without APCs (OpenAlex, n.d.). Together, these categories provide a useful framework for examining how patterns of retractions intersect with different publishing models.

Examining retractions by access status, Figure 2 shows the distribution of retracted research across access types. Closed access publications account for the largest number of retractions, emphasizing that subscription-based, paywalled journals still remain the dominant form of scholarly publishing. However, a substantial proportion of retractions are associated with OA models, particularly gold OA, which represents the largest number of retractions among OA modalities.

'a substantial proportion of retractions are associated with OA models, particularly gold OA'

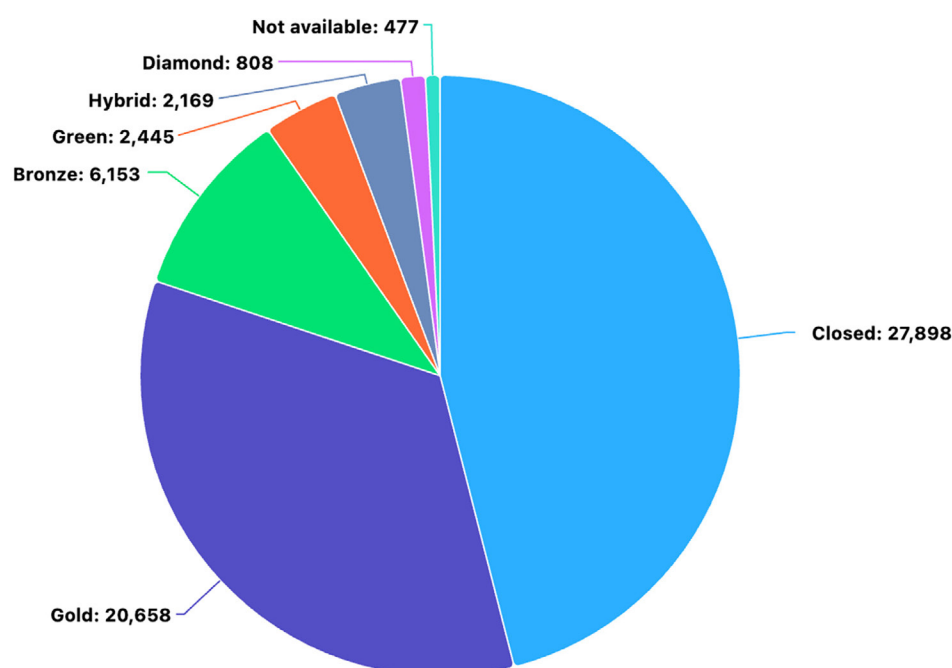


Figure 2. Retractions by access type, 2009–2024

Bronze open access also accounts for a notable number of retractions (6,153), exceeding the number of retractions for green, hybrid and diamond OA combined. This is significant given that bronze OA articles are often free to read but lack standardized licensing and clear reuse rights, potentially complicating post-publication oversight and correction workflows. Green and hybrid OA articles show more moderate retraction counts (2,445 and 2,169), while diamond OA retractions represent a comparatively small number (808), consistent with its more limited scale within the broader publishing ecosystem.

The distribution of retractions by access type highlights that while retractions remain most common in closed access venues, a meaningful and growing portion occurs within OA modalities, especially gold OA. This pattern demonstrates the importance of examining how publishing models, economic incentives and access pathways shape not only the visibility of

7 research but also the processes through which errors, misconduct and integrity breaches are identified and corrected.

While raw retraction counts by access type provide a useful picture of where retractions are concentrated, they do not account for differences in publication volume across access types. To contextualize these counts, retraction rates were calculated by dividing the total retractions for each modality by the total number of journal articles published for that modality using data from OpenAlex (Table 1). From 2009 to 2024, closed access publications accounted for 27,898 retractions out of 44,598,966 total publications, representing a retraction rate of 0.063%. Gold OA accounted for 20,658 retractions out of 6,728,889 total publications (0.307%), hybrid OA accounted for 2,169 retractions out of 4,098,866 total publications (0.053%), green OA accounted for 2,445 retractions out of 5,220,205 total publications (0.047%), bronze OA accounted for 6,153 retractions out of 5,868,629 total publications (0.105%) and diamond OA accounted for 808 retractions out of 11,910,644 total publications (0.007%).

Modality	Retractions	Total Publications	Retraction Rate
Closed	27,898	44,598,966	0.063%
Gold	20,658	6,728,889	0.307%
Hybrid	2,169	4,098,866	0.053%
Green	2,445	5,220,205	0.047%
Bronze	6,153	5,868,629	0.105%
Diamond	808	11,910,644	0.007%

Table 1. Retractions and retraction rates by access type, 2009–2024

These proportional figures offer a more meaningful basis for comparison than raw counts alone. Despite accounting for fewer raw retractions than closed access publications, gold OA articles were retracted at a rate nearly five times higher than closed access articles (0.307% vs. 0.063%), suggesting that the volume dominance of closed access retractions masks a substantially elevated integrity risk within APC-funded gold OA publishing. Hybrid OA, by contrast, shows a retraction rate (0.053%) slightly lower than closed access, complicating any simple characterization of OA as inherently more prone to integrity failures and pointing instead to the specific role of incentive structures within high-volume gold OA models. The most striking comparison is between gold OA and diamond OA. Diamond OA, where no APCs are charged and publisher revenue is not tied to acceptance volume, shows a retraction rate of just 0.007% – more than 40 times lower than gold OA. This comparison directly isolates the potential influence of APC incentive structures from OA status itself, and the magnitude of the difference is consistent with the hypothesis that revenue models tied to publication volume create meaningful structural pressures on editorial quality.

'Diamond OA, where no APCs are charged ... shows a retraction rate of just 0.007%'

### Retractions by access status over time

Although retractions of closed access articles accounted for a significant portion of retractions overall, there has been a noticeable shift in the retraction of OA articles in recent years. Figure 3 shows how the access status of retracted articles has shifted over time, highlighting a clear change in the composition of retractions as OA publishing has continued to expand. Early on, particularly from 2009 to 2011, retractions were overwhelmingly associated with closed access articles, which accounted for nearly all retracted articles during this period. At that time, open access retractions (gold, hybrid, green and diamond) represented only a small fraction of the total.

This pattern began to change around 2012, as the share of retractions involving OA articles began to steadily increase. The most pronounced change in OA article retractions from 2009 to 2024 was with gold OA. Shifting from a relatively modest share of retractions in the early 2010s, gold OA retractions grew consistently across the decade. By the 2020s, gold OA accounted for a substantial, and in some years (2023 and 2024) a majority, portion of all retractions, which is a substantial reversal from the initial prevalence of closed access retractions.

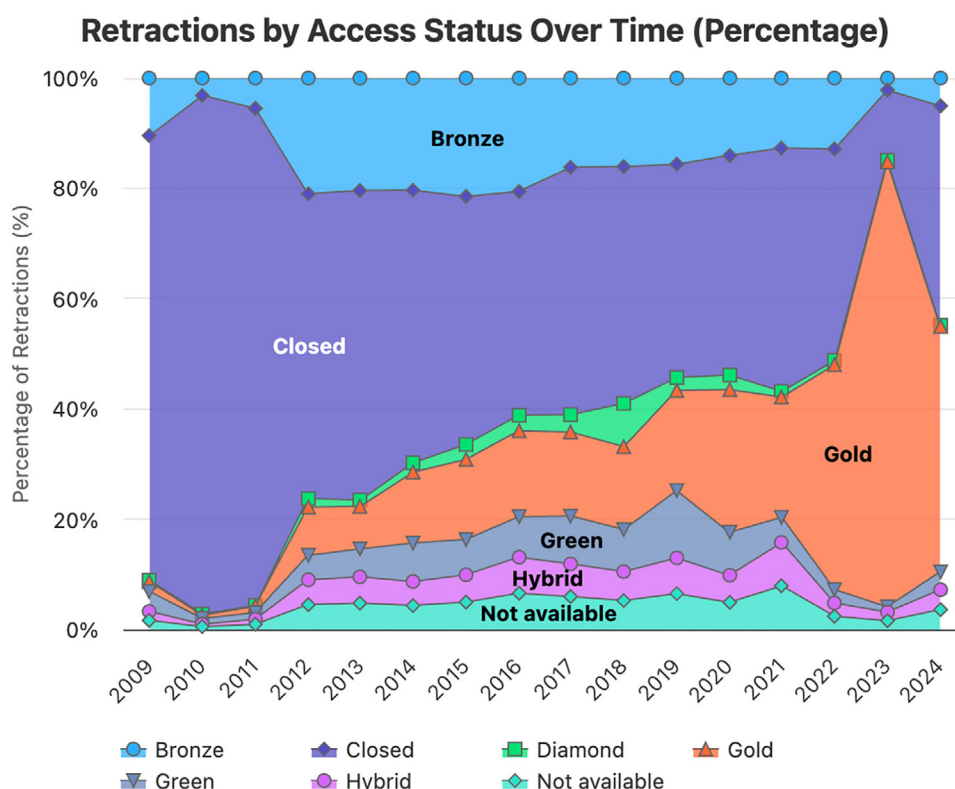


Figure 3. Retractions by access type over time, 2009–2024

Other OA modalities (hybrid, green, bronze and diamond) showed more modest and uneven growth in the number of retractions. The number of hybrid and green OA retractions remained relatively low, fluctuating over time but never coming remotely close to approaching the scale of gold OA retractions. Outside of gold OA, bronze OA accounted for the largest share of retractions, while diamond OA represented a small fraction of retractions and in some years seemed to disappear almost entirely. For a small proportion of retractions, the access type was unavailable through OpenAlex, and these retractions fall under the category of 'Not available'.

While closed access research continues to account for a large portion of retractions, the growing prominence of gold OA among retracted publications mirrors the broader expansion of APC-funded publishing. This shift raises important questions about how editorial practices, economic incentives and post-publication oversight operate within different access models, and it suggests that research integrity discussions must account for the governance structures of all publishing modalities, OA included.

'the growing prominence of gold OA among retracted publications mirrors the broader expansion of APC-funded publishing'

### Retractions by subject domain – gold and hybrid OA

In addition to examining the access types of retracted research, it is equally important to consider how retractions are distributed across subject domains, particularly within the two most prevalent OA modalities: gold and hybrid. Within the Retraction Watch dataset, retracted articles are frequently assigned to multiple subject domains; as a result, the counts reported for each domain reflect overlapping classifications rather than unique publications.

The distribution of retractions by subject domain provides additional insight into where research integrity issues most frequently emerge across the scholarly publishing landscape. Figure 4 shows that the majority of gold and hybrid OA retractions are heavily concentrated within three subject domains: Health Sciences, Basic Life Sciences and Business and Technology. Health Sciences accounts for the most OA retractions, with 9,024 retractions across gold and hybrid OA, followed by Basic Life Sciences (8,888) and Business and Technology (8,750).

'OA retractions are heavily concentrated within three subject domains: Health Sciences, Basic Life Sciences and Business and Technology'

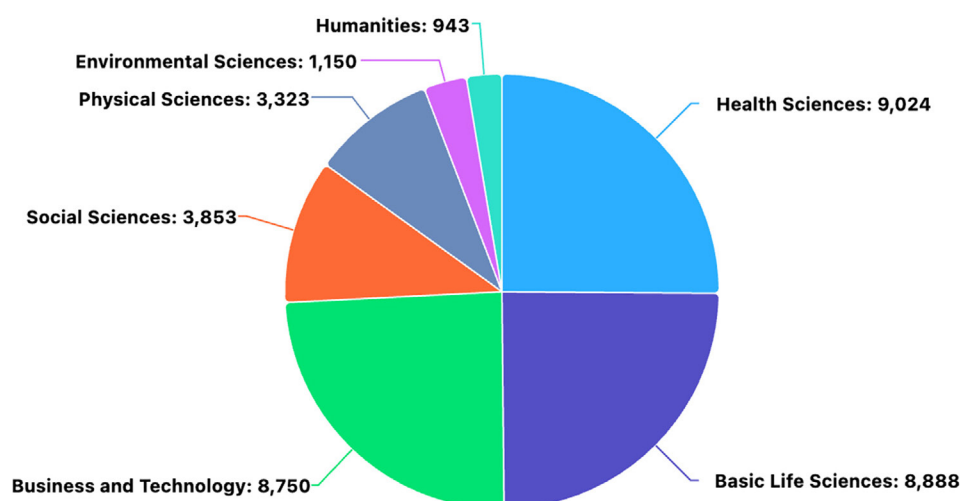


Figure 4. OA Retractions by subject domain, 2009-2024

Together, these three areas account for the bulk of retracted research in the dataset, reflecting where much of the scholarly output, and scrutiny, is concentrated. Retractions are notably less common in Social Sciences (3,853) and Physical Sciences (3,323), and even lower gold and hybrid OA retraction rates can be seen in Environmental Sciences (1,150) and Humanities (943), which account for relatively few retractions. Overall, the pattern underscores how retractions are far more prevalent in subject domains with high publication volume and rapid research cycles, while they remain comparatively small in lower-volume publishing subject domains.

### Retractions by country – gold and hybrid OA

Looking at the distribution of OA retractions by country helps move beyond raw retraction counts to a better understanding of where retraction activity is most visible. The ten countries with the most OA retractions were China, India, the United States, Saudi Arabia, Pakistan, South Korea, Russia, Ethiopia, the United Kingdom and Iran.

China accounts for the largest number of OA retractions by far (15,653), exceeding the combined totals of all other listed countries. India has the second-highest number, with 1,749 OA retractions, followed by the United States (1,586) and Saudi Arabia (1,153). The remaining countries (Pakistan, South Korea, Russia, Ethiopia, the United Kingdom and Iran) each have several hundred OA retractions.

Importantly, these raw counts should not be read as retraction rates. To provide more meaningful context, [Table 2](#) includes each country's estimated total number of OA journal articles published and the corresponding retraction rate as a proportion of that output. These proportional figures reveal a substantially different picture than raw counts alone. Ethiopia, which ranks eighth by raw retraction count, has the highest retraction rate among the ten countries at 0.604%, followed closely by China at 0.547% and Saudi Arabia at 0.446%. By contrast, the United States and the United Kingdom, which appear in the top ten largely by virtue of their high publication volume, show retraction rates of just 0.032% and 0.026%, respectively, which are among the lowest in the group. These differences suggest that raw retraction counts are heavily shaped by overall research output and that countries with large publishing outputs can accumulate high retraction counts while maintaining comparatively low retraction rates. Conversely, countries with smaller overall output but elevated retraction rates, such as Ethiopia and Saudi Arabia, may reflect the influence of institutional incentive structures that prioritize publication volume over research quality, the rapid expansion of APC-funded publishing with limited oversight infrastructure or gaps in peer-review capacity ([Hassen, 2024](#); [Sawahel, 2025](#)). As with the modality-level analysis, these figures should be interpreted with caution given the limitations in OpenAlex coverage consistency across countries and over time, and they are best understood as a starting point for further investigation rather than definitive measures of national research integrity.

'retractions are far more prevalent in subject domains with high publication volume and rapid research cycles'

Country	OA Publications	OA Retractions	OA Retraction %
China	2,860,214	15,653	0.547%
India	1,250,799	1,749	0.140%
United States	4,951,758	1,586	0.032%
Saudi Arabia	258,260	1,153	0.446%
Pakistan	234,218	674	0.288%
South Korea	658,017	504	0.077%
Russia	802,626	469	0.058%
Ethiopia	73,964	447	0.604%
United Kingdom	1,688,234	446	0.026%
Iran	426,009	417	0.098%

Table 2. OA retractions, total OA publications and retraction rates by country (top ten by raw retraction count), 2009–2024

### Retractions by reason – gold and hybrid OA

An analysis of the reasons for retractions among gold and hybrid OA shows that they largely mirror the broader patterns of retractions overall, regardless of access status. Issues tied to oversight and research reliability were the main reasons for many retractions. As outlined in [Table 3](#), journal or publisher investigations were the most common trigger, accounting for 14,844 OA retractions, followed closely by unreliable results (12,064) and third-party investigations (11,972). These same issues also dominate retractions across scholarly publishing as a whole, with journal or publisher investigations accounting for 24,120 retractions, unreliable results accounting for 15,795 retractions and third-party investigations accounting for 14,683 retractions, underscoring that OA articles are not being retracted for fundamentally different reasons but due to many of the same breakdowns in research and editorial processes seen elsewhere.

‘Issues tied to oversight and research reliability were the main reasons for many retractions’

Reason	Total OA Retractions	Gold OA Retractions	Hybrid OA Retractions
Investigation by Journal/Publisher	14,844	14,210	634
Unreliable Results	12,064	11,874	190
Investigation by Third Party	11,972	11,765	207
Concerns/Issues About Data	10,527	10,199	328
Concerns/Issues about Referencing/Attributions	9,068	8,967	101
Concerns/Issues with Peer Review	8,970	8,919	51
Paper Mill	8,935	8,763	172
Concerns/Issues About Results	7,001	6,890	111
Randomly Generated Content	6,681	6,514	167
Fake Peer Review	4,035	3,696	339

Table 3. Reasons for OA retractions (top ten), including gold and hybrid OA retractions

Across all reasons, gold OA retractions account for more than 91% of all retractions. Referring to the reasons in [Table 3](#), gold OA accounts for 95.7% of retractions due to journal or publisher investigations (14,844 total OA retractions); 98.4% of retractions due to unreliable results (12,064 total OA retractions); 98.2% of retractions due to investigations by a third party (11,972 total OA retractions); 96.8% of retractions due to concerns/issues about data (10,527 total OA retractions); 98.9% of retractions due to concerns/issues about referencing/attribution (9,068 total OA retractions); 99.4% of retractions due to concerns/issues with peer review (8,970 total OA retractions); 98% of retractions due to paper mill activity (8,935 total OA retractions); 98.4% of retractions due to concerns/issues about results (7,001 total OA retractions);

‘gold OA accounts for 95.7% of retractions due to journal or publisher investigations’

11 97.5% of retractions due to randomly generated content (6,681 total OA retractions); and 91.6% of retractions due to fake peer review (4,035 total OA retractions). Rather than indicating unique or isolated problems within OA, the data reinforce how OA publishing is embedded within the same research integrity challenges facing scholarly publishing as a whole.

### Retractions by publisher – gold and hybrid OA

Examining OA retractions at the publisher level provides insight into how OA retraction activity is distributed across different publishers and OA publishing models. Publishers with the most OA retractions include Hindawi, Springer Nature, IOP Publishing, PLOS, Elsevier, Taylor & Francis, Frontiers, Wiley, Spandidos and MDPI (Table 4).

Publisher	Total OA Retractions	Gold OA Retractions	Hybrid OA Retractions
Hindawi	11,432	11,432	0
Springer Nature	2,198	1,675	523
IOP Publishing	1,159	1,152	7
PLOS	1,137	1,137	0
Elsevier	973	743	230
Taylor & Francis	932	867	65
Frontiers	435	435	0
Wiley	397	318	79
Spandidos	350	197	153
MDPI	340	340	0

Table 4. Total OA retractions by publisher (top ten), including gold and hybrid OA retractions

Hindawi, a gold OA publisher acquired by Wiley in 2021, accounted for 50% of all OA retractions, with 11,432 retractions. As a result of widespread concerns about research integrity issues, such as paper mills, Wiley ceased using the Hindawi brand name in 2023 following a US\$18 million revenue decline (Kincaid, 2023). Other exclusively gold OA publishers with the highest number of retractions were PLOS (1,137), Frontiers (435) and MDPI (340).

Even among publishers with mixed OA portfolios, gold OA retractions accounted for the majority of all retractions. Of Springer Nature’s 2,198 OA retractions, gold OA accounted for 76% (1,675). The same is true for other large commercial publishers like Elsevier (76% gold OA), Wiley (80.1% gold OA) and Taylor & Francis (93% gold OA). For IOP Publishing, gold OA accounted for nearly all of its retractions (99.4%), while the OA retractions for Spandidos were more evenly split, with gold OA accounting for 56% and hybrid OA accounting for 44% of retractions. The concentration of retractions in gold OA journals highlights the need for closer and continued scrutiny of APC-based publishing models, particularly where revenue is closely tied to article volume, as this incentivizes publishing more to increase revenue, even at the expense of quality.

### APCs by publisher – gold and hybrid OA

Based on available data from OpenAlex, in addition to the total number of OA retractions, there was a substantial sum of APCs paid to publishers for research that was later retracted. In total, US\$41,962,729 was paid in APCs for retracted research. Gold APCs accounted for US\$37,965,478, while APCs for hybrid OA accounted for US\$3,997,251. A breakdown of APCs by publisher (Table 5) shows that the top ten publishers accounted for US\$38,930,489.60 (92.7%) of APCs. Of those publishers, Hindawi accounted for more than half, with over US\$23,284,456 in APCs.

Including Hindawi, there were seven publishers that accounted for more than US\$1 million in APCs, including Springer Nature (US\$4,707,610.21), Taylor & Francis (US\$2,411,102.50), PLOS (US\$2,187,288.46), Elsevier (US\$1,990,311.62), Frontiers (US\$1,221,645.69) and Wiley (US\$1,013,067.51).

‘there was a substantial sum of APCs paid to publishers for research that was later retracted’

Publisher	Total APCs	Gold OA APCs	Hybrid OA APCs
Hindawi	\$23,284,456	\$23,284,456	\$0
Springer Nature	\$4,707,610.21	\$3,345,861.12	\$1,361,749.09
Taylor & Francis	\$2,411,102.50	\$2,356,219.03	\$54,883.47
PLOS	\$2,187,288.46	\$2,187,288.46	\$0
Elsevier	\$1,990,311.62	\$1,232,628.03	\$757,683.59
Frontiers	\$1,221,645.69	\$1,221,645.69	\$0
Wiley	\$1,013,067.51	\$744,342.86	\$268,724.65
American Society for Biochemistry and Molecular Biology (ASBMB)	\$781,210	\$5,310	\$775,900
MDPI	\$779,117.61	\$779,117.61	\$0
Portland Press	\$554,680	\$554,680	\$0

Table 5. Total APCs by publisher (top ten), including gold and hybrid APCs, in US dollars

Across most publishers in the top ten, gold OA accounted for the vast majority of APCs. The notable exception to this was the American Society for Biochemistry and Molecular Biology, where nearly all APCs (US\$775,900 out of US\$781,210) were for hybrid OA articles published prior to their transition to a gold OA model in 2021 (Hart, 2020). Overall, APCs for retracted research were heavily concentrated among these ten publishers and were overwhelmingly tied to gold OA publishing rather than hybrid models.

## Discussion

The expansion of OA publishing, particularly APC-based OA publishing, has reshaped how research is shared, and this expansion has also created financial incentive structures that govern editorial decision-making. While OA models have expanded access to research, APC-based models, especially gold OA, have also directly tied publisher revenue to publication volume. In doing so, they create conditions where increasing acceptance rates, and thereby increasing publication output, is financially advantageous, even as expectations around quality, peer review and editorial oversight grow more difficult to sustain at scale.

This tension is most visible among large commercial publishers (Elsevier, Springer Nature and Wiley) and exclusively gold OA publishers (Frontiers, MDPI and PLOS), all of which generate a significant amount of revenue from APCs. From 2019 to 2023, these publishers generated more than US\$8 billion in revenue from APCs (Butler et al., 2024).

Additionally, in investor communications from large commercial publishers, OA growth is frequently framed in terms of rising submissions, expanding journal portfolios and improved (increased) acceptance rates – metrics that signal financial health but also reveal how closely editorial decision-making and revenue are linked. For example, Taylor & Francis has explicitly discussed higher publication volume and pay-to-publish revenues as part of its growth narrative, positioning it as being responsive to author demand and market expansion. It highlighted that the ‘increasing volumes of research publication and high demand for open access content...are fueling the growth in the pay-to-publish revenues with a revenue growth trend of over 20%’ (Informa, 2025). While increased publication volume does not inherently indicate weaker quality control, it does raise questions about how editorial thresholds are negotiated and decided when journal performance is evaluated through revenue and volume metrics. Additionally, Springer Nature highlighted strong article growth (10%) overall, outpacing the market as a whole. Article growth for the full OA (gold OA) model was well over 25% (Springer Nature, 2025a).

‘OA growth ... metrics that signal financial health but also reveal how closely editorial decision-making and revenue are linked’

At the far end of APC-based OA publishing growth are megajournals such as *Scientific Reports*, which highlight the revenue potential of high-volume gold OA models. With annual output approaching 45,000 articles in 2025, APCs translate into tens of millions of dollars in revenue, with plausible estimates exceeding \$100 million (Butcher, 2025). At this scale, any breakdowns in peer review or editorial oversight, whether systemic or irregular, can

13 have wide-ranging effects, both financially and reputationally. This has been the case with other megajournals such as *Heliyon* (Elsevier) and *Cureus* (Springer Nature) – for example, Clarivate’s Web of Science (WoS) paused indexing new content from these journals and later delisted *Cureus* from WoS due to research integrity concerns (Joelving, 2024).

These issues and financial incentives become even more interesting when considering how publishers handle APCs for retracted articles. While many publishers have their own policies that may vary slightly, in the cases of major publishers like Elsevier, Wiley, Springer Nature and Taylor & Francis, APCs are not refunded when articles are retracted due to misconduct, data reliability issues or compromised peer review. Elsevier’s policy states, ‘No refund or credit will be offered in the following circumstances: Article retraction or removal: Elsevier has provided publishing services. The later retraction or removal of the article is typically for reasons beyond our control, and does not detract from the publishing services provided, nor from our ongoing maintenance of the scientific record, e.g., corrections to the record’ (Elsevier, n.d.). While there are some circumstances where Elsevier will refund APCs, such as incorrect licensing or a delay in providing OA (Elsevier, n.d.), Wiley does not offer APC refunds for any reason (Wiley, n.d.). Like Elsevier, Springer Nature will refund APCs if an error has occurred on its part but does not offer APC refunds for retracted research (Springer Nature, 2025b). Similarly, Taylor & Francis does not refund APCs for retracted research. Its policy states, ‘After publication, we may retract or remove the article from publication if we reasonably consider this is necessary to ensure the integrity of the scholarly record. In [that] situation, we shall be entitled to retain any Article Publishing Charge you have paid’ (Taylor & Francis, n.d.). These policies enable publishers to retain revenue from articles that are later retracted. However, there are some publishers that have experimented with alternative approaches. For example, IOP Publishing has committed to donating APCs from retracted articles to Research4Life (Trager, 2024). While such measures do not resolve the deeper incentive misalignments of APC-based publishing, they do signal the growing recognition that retractions raise ethical questions about authorship as well as publisher responsibility.

‘in the cases of major publishers ... APCs are not refunded when articles are retracted due to misconduct, data reliability issues or compromised peer review’

These financial and policy dynamics intersect with entrenched citation cultures and narratives surrounding the perception of the open access citation advantage (OACA), as well as the broader publish-or-perish environment that shapes and influences academic careers. In research environments where publication counts and citations function as primary signals of productivity and impact, the perception that OA confers a citation advantage can incentivize publishing decisions that prioritize output over rigor. Some scholars argue that APC-based OA, specifically gold OA, is leading to research integrity issues and corruption while also incentivizing publishers to publish low-interest and low-quality research (Morgan and Smaldino, 2025). These pressures are further intensified within APC-based publishing models, where publisher revenue is directly linked to article volume and where increased research output aligns with – and is required for – both institutional performance metrics and individual career incentives. Under these conditions, citation counts and h-indexes may take precedence over careful scrutiny of research quality, allowing integrity concerns to be overlooked, particularly when retracted or otherwise questionable articles continue to be cited and remain embedded within the scholarly communication ecosystem. For APC-based publishing, this essentially results in a scenario where authors are not just paying for article processing charges; they are paying for citations.

## Conclusion

These patterns suggest that research integrity challenges associated with OA publishing, particularly gold OA, are not simply the result of individual misconduct or isolated editorial lapses. Rather, they reflect deeper structural tensions within APC-based publishing and research systems, where financial incentives, evaluation metrics and scholarly norms are closely intertwined. Addressing these challenges will require co-ordinated efforts among

‘research integrity challenges associated with OA publishing, particularly gold OA, are not simply the result of individual misconduct or isolated editorial lapses’

## 14 publishers, funders, libraries and research institutions to better align OA, and openness in general, with quality, accountability and sustained trust.

### Data accessibility statement

The data for this study can be found at <https://zenodo.org/records/15185273>.

### Abbreviations and Acronyms

A list of the abbreviations and acronyms used in this and other *Insights* articles can be accessed here – click on the following URL and then select the ‘full list of industry A&As’ link: <http://www.uksg.org/publications#aa>.

### Competing interests

The author has declared no competing interests.

### References

- Björk, B., & Solomon, D. (2012). Open access versus subscription journals: A comparison of scientific impact. *BMC Medicine*, 10(1), 73. <https://doi.org/10.1186/1741-7015-10-73> (accessed 8 June 2026).
- Bordignon, F. (2025). Moving open repositories out of the blind spot of initiatives to correct the scholarly record. *Learned Publishing*, 38(2), e1655. <https://doi.org/10.1002/leap.1655> (accessed 8 June 2026).
- Brainard, J. (2026). Some guest editors pack special issues with their own articles. *Science*, 391(6783), 332–333. <https://doi.org/10.1126/science.aef6706> (accessed 8 June 2026).
- Butcher, J. (2025, December 3). *Another one bites the dust*. Journalology. <https://newsletter.journalology.com/p/another-one-bites-the-dust> (accessed 8 June 2026).
- Butler, L.-A., Hare, M., Schönfelder, N., Schares, E., Alperin, J. P., & Haustein, S. (2024). An open dataset of article processing charges from six large scholarly publishers (2019–2023). (Preprint). *arXiv*. <https://doi.org/10.48550/arXiv.2406.08356> (accessed 08 June 2026).
- Chen, X., Bharti, N., & Marsteller, M. (2021). Use of bibliometrics data to understand the citation advantages of different open access categories in covid-19 related studies. *Proceedings of the Association for Information Science and Technology*, 58(1), 410–414. <https://doi.org/10.1002/pr2.469> (accessed 8 June 2026).
- Ellingson, M. K., Shi, X., Skydel, J. J., Nyhan, K., Lehman, R., Ross, J., S, Wallach., & J, D. (2021). Publishing at any cost: A cross-sectional study of the amount that medical researchers spend on open access publishing each year. *BMJ Open*, 11(2), 1–10. <https://bmjopen.bmj.com/content/11/2/e047107> (accessed 8 June 2026).
- Elsevier. (n.d). *Pricing*. Elsevier Policies. <https://www.elsevier.com/about/policies-and-standards/pricing> (accessed 8 June 2026).
- Gasparian, A. Y., Ayyvazyan, L., & Kitas, G. D. (2013). Open access: Changing global science publishing. *Croatian medical journal*, 54(4), 403–406. <https://doi.org/10.3325/cmj.2013.54.403> (accessed 8 June 2026).
- Hart, G. (2020, May 13). *ASBMB journals are going open access*. American Society for Biochemistry and Molecular Biology. <https://www.asbmb.org/asbmb-today/opinions/051320/asbmb-journals-are-going-open-access> (accessed 8 June 2026).
- Hassen, A. E. (2024, March 13). *Intellectual harvest — Ethiopia’s research output [by area of study]*. Reqiq. <https://reqiq.co/intellectual-harvest-ethiopias-research-output-by-area-of-study/> (accessed 8 June 2026).
- Informa. (2025, November 17). *Informa 2025 Capital Markets Day transcript*. Author. <https://www.informa.com/globalassets/documents/investor-relations/2025/informa-2025-capital-markets-day-transcript.pdf> (accessed 8 June 2026).
- Jacobs, P. (2025, December 12). *NIH’s proposed caps on open-access publishing fees roil scientific community*. Science. <https://www.science.org/content/article/nih-s-proposed-caps-open-access-publishing-fees-roil-scientific-community> (accessed 8 June 2026).
- Jahn, N., & Tullney, M. (2016). A study of institutional spending on open access publication fees in Germany. *PeerJ*, 4, e2323. <https://doi.org/10.7717/peerj.2323> (accessed 8 June 2026).
- Joelving, F. (2024, September 30). *Web of Science puts mega-journals Cureus and Heliyon on hold*. Retraction Watch. <https://retractionwatch.com/2024/09/30/web-of-science-puts-mega-journals-cureus-and-heliyon-on-hold/#more-130124> (accessed 8 June 2026).
- Khoo, S. Y.-S. (2019). Article processing charge hyperinflation and price insensitivity: An open access sequel to the serials crisis. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, 29(1), 1–18. <https://doi.org/10.18352/lq.10280> (accessed 8 June 2026).
- Kincaid, E. (2023, December 6). *Wiley to stop using “Hindawi” name amid \$18 million revenue decline*. Retraction Watch. <https://retractionwatch.com/2023/12/06/wiley-to-stop-using-hindawi-name-amid-18-million-revenue-decline/> (accessed 8 June 2026).
- Morgan, T. J. H., & Smaldino, P. E. (2025). Author-paid publication fees corrupt science and should be abandoned. *Science & Public Policy*, 52(5), 805–809. <https://doi.org/10.1093/scipol/scaf026> (accessed 8 June 2026).
- OpenAlex. (n.d). *Open access (OA)*. Author. <https://help.openalex.org/hc/en-us/articles/24347035046295-Open-Access-OA> (accessed 8 June 2026).
- Retraction Watch. (2025). *Retraction watch data*. Github. <https://gitlab.com/crossref/retraction-watch-data> (accessed 8 June 2026).
- Retraction Watch. (n.d). *Retraction Watch Database User Guide Appendix A: Fields*. Retraction Watch. <https://retractionwatch.com/retraction-watch-database-user-guide/retraction-watch-database-user-guide-appendix-a-fields/> (accessed 8 June 2026).
- Sawahel, W. (2025, January 17). *Research output soars by 320%, but what are metrics hiding?* University World News. <https://www.universityworldnews.com/post.php?story=2025011711243560> (accessed 8 June 2026).

- Siler, K., Hausteijn, S., Smith, E., Larivière, V., & Alperin, J. P. (2018). Authorial and institutional stratification in open access publishing: The case of global health research. *PeerJ*, 6, e4269. <https://doi.org/10.7717/peerj.4269> (accessed 8 June 2026).
- Solomon, D., & Björk, B. (2016). Article processing charges for open access publication—the situation for research intensive universities in the USA and Canada. *PeerJ*, 4(e2264). <https://doi.org/10.7717/peerj.2264> (accessed 8 June 2026).
- Springer Nature. (2025a, November 12). *SN Presentation Q3 2025 final webcast*. Springer Nature Investor Relations. [https://www.ir.springernature.com/media/document/ce9b4db1-0aae-487e-a1fd-4d97143deb7e/assets/SN\\_Presentation\\_Q3\\_2025\\_final\\_webcast.pdf](https://www.ir.springernature.com/media/document/ce9b4db1-0aae-487e-a1fd-4d97143deb7e/assets/SN_Presentation_Q3_2025_final_webcast.pdf) (accessed 8 June 2026).
- Springer Nature. (2025b, February 10). *APC refund policy*. Author. <https://support.springernature.com/en/support/solutions/articles/6000227581-apc-refund-policy> (accessed 8 June 2026).
- Taylor & Francis. (n.d). *Open access frequently asked questions*. Author. <https://www.tandfonline.com/openaccess/faqs> (accessed 8 June 2026).
- Tocco, E. G., Mayhew, M. M., Mercante, M.G., Kuchimanchi, N., Moheb, M. E., Shen, C., Tsung, A., & Witt, R. G. (2025). Variability in article processing charges of open-access publishing across medical specialties. *Plos One*, 20(7), e0320684. <https://doi.org/10.1371/journal.pone.0320684> (accessed 8 June 2026).
- Trager, R. (2024, January 8). *Institute of physics donates authors fees from retracted papers to publishing charity*. Chemistry World. <https://www.chemistryworld.com/news/institute-of-physics-donates-authors-fees-from-retracted-papers-to-publishing-charity/4018725.article> (accessed 8 June 2026).
- UK Research Integrity Office. (2026). *What is research integrity?*. Author. <https://ukrio.org/research-integrity/what-is-research-integrity/> (accessed 8 June 2026).
- Vervoort, D., Luc, J. G. Y., Sá, M. P. B. O., & Etchill, E. W. (2021). Open Access and Article Processing Charges in Cardiology and Cardiac Surgery Journals: A Cross-Sectional Analysis. *Brazilian Journal of Cardiovascular Surgery*, 36(4), 453–460. <https://pubmed.ncbi.nlm.nih.gov/34617426/> (accessed 18 June 2026).
- Wiley. (n.d). *Article processing charges (APCs)*. Wiley Author Services. <https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/article-publication-charges/index.html> (accessed 8 June 2026).
- Zheng, E., & Fu, H. (2024). A comparative study on characteristics of retracted publications across different open access levels. *Journal of Data and Information Science*, 9(2), 22–40. <https://doi.org/10.2478/jdis-2024-0010> (accessed 8 June 2026).

**Article copyright: © 2026 Ben Rawlins. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and distribution provided the original author and source are credited.**



Corresponding author:  
Ben Rawlins  
Associate Dean for Outreach, Engagement and Collections  
University of Kentucky  
United States  
E-mail: [ben.rawlins@uky.edu](mailto:ben.rawlins@uky.edu)  
ORCID ID: <https://orcid.org/0000-0002-8277-803X>

To cite this article:  
Rawlins B, "Research integrity and open access models: insights from Retraction Watch and OpenAlex," *Insights*, 2026, 39: 14, 1–15; DOI: <https://doi.org/10.1629/uksg.763>

Submitted on 20 January 2026

Accepted on 23 March 2026

Published on 7 July 2026

Published by UKSG in association with Ubiquity Press.