Tracking the Data Quality Landscape of Retracted Papers: Flag Usage in Titles and Changes in DOI Retraction Status

Luyang Si[*1], Malik Oyewale Salami[*], and Jodi Schneider[2**]

[*] *ls36@illinois.edu & sly3022123556@gmail.com*, *malikos2@illinois.edu*
0009-0008-7883-4757; 0000-0002-2329-5660
School of Information Sciences, University of Illinois Urbana-Champaign, USA

[**] *jodi@wisc.edu* & *jschneider@pobox.com*
0000-0002-5098-5667
Information School, University of Wisconsin–Madison, USA

Abstract Inconsistent retraction indexing can perpetuate the spread of unreliable findings. We investigate the extent of errors in retraction indexing and the reliability of title flags, such as *retracted articles*, *removal*, etc. We manually checked the retraction status on publisher websites for 925 DOIs that Crossref indexed as retracted as of April 2023. We sampled DOIs with title flags that may indicate retracted status (e.g., *retracted article:*) (n=331) and retraction notice status (e.g., *removal notice)* (n=594). According to publisher websites, 661/925 DOIs (71.46%) in our sample are retraction notices. However, these same DOIs are indexed as retracted papers in Crossref. As of June 2025, all but one of the 585 DOIs with retraction notice flags were indeed retraction notices, whereas some DOIs have retracted paper flags (77/195, 39.49%). Our findings highlight significant errors in Crossref's retraction indexing despite improvements in data quality as retraction indexing changed over time.

Keywords: retracted publications, retraction indexing, data quality, metadata errors

1. Introduction

Retracted papers are often cited inadvertently, in part because researchers have trouble identifying which papers have been retracted. For example, Minetto et al. (2024) found that 78.7% of surveyed authors were unaware of the retraction status of the articles they cited, most commonly due to inadequate notification in research databases. Similarly, De Cassai et al. (2022) reported that 89% of authors who cited retracted literature in anesthesiology were unaware that the cited articles had been retracted despite many of the authors checking references for retractions. The Committee on Publication Ethics (COPE Council, 2025) recommends that retracted publications be clearly and consistently labeled across all platforms, including journal webpages and databases. Each retracted publication should be explicitly marked in the paper title, such as by prepending "Retracted:" to the title (COPE Council, 2025).

Authors' lack of awareness of citing retracted publications points to broader challenges in the scholarly communication ecosystem. Publisher websites are known to have inaccuracies, failing to display retraction status for known retracted items in past studies (Bakker & Riegelman, 2018;

[1] Work completed as an undergraduate student at University of Illinois Urbana-Champaign.
[2] Work started at University of Illinois Urbana-Champaign, continued at the Harvard Radcliffe Institute for Advanced Study, Cambridge, MA, USA, and at University of Wisconsin–Madison.

Bakker et al., 2024; Yan et al., 2016), as well as inconsistencies in how retractions are labeled even across the same publisher website (Suelzer et al., 2021).

Two databases cannot both be correct when they disagree on a publication's retraction status. This disagreement between databases indicates that one has a mistake. Such mistakes seem to be more common from 2023-2024 (Salami et al., 2024), when an overwhelming number of papers were retracted (Van Noorden, 2023). This inconsistent retraction indexing may hinder authors' ability to identify retracted work and can perpetuate the spread of unreliable findings, increasing the likelihood that such findings are cited or reused.

We found that databases disagree in our previous study (Schneider et al., 2023). In particular, we found that a paper indexed as retracted in one database may be covered in another database, but not indexed as retracted. This led to our union list of purportedly retracted publications (Schneider et al., 2023). We observed systematic metadata inconsistencies in the titles of papers in our union list. For example, instead of standardized retraction flags (e.g., "RETRACTED" or "RETRACTED ARTICLE") as suggested by the COPE Council (2025), some papers' titles included phrases suggesting that they were retraction notices, such as "RETRACTION" or "RETRACTION NOTICE". This mismatch suggests the possibility of data quality problems which cannot be resolved without direct case-by-case manual examination.

Prior work on data quality leveraged links between retraction notices and retracted publications (Schmidt, 2018) or examined paper titles (Uppala et al., 2022). This paper uses a combined approach: examining paper titles and changes in what is indexed as retracted.

2. Background

2.1 Definitions

"**Retraction** is a mechanism for correcting the literature and alerting readers to errors or inaccurate data that articles may have, such as seriously flawed or erroneous content or data that their findings and conclusions cannot be relied upon." (COPE Council, 2025, p. 2).

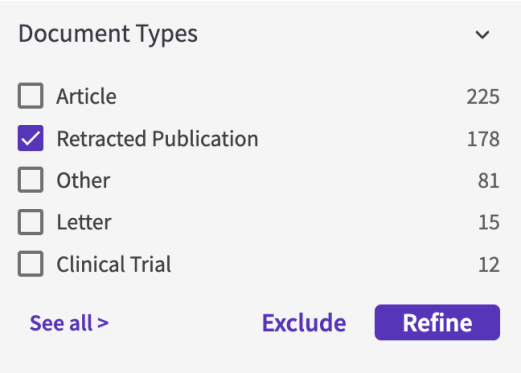
A **Retracted paper**, or **Retracted publication**, is a paper officially retracted by the publisher for flaws or errors in its content or data.

A **Retraction notice** is an editorial notice indicating that an item is retracted. It may state the reasons for the retraction (COPE Council, 2025).

Retraction status distinguishes retracted papers from retraction notices and from normal articles.

Retraction indexing enables database users to systematically retrieve and identify retracted papers easily. Figure 1 shows an example of retracted indexing, which enables users to filter document types.

Figure 1: The "Retracted Publication" category on Web of Science is an example of retraction indexing.



A **Title flag** (or **flag**) refers to a specific phrase that helps researchers guess whether an item is a retracted paper or a retraction notice. A flag may appear at the beginning or end of a title, or may constitute the full title. Examples of flags are shown in Tables 1 and 2 in variant formats. Besides variant flag formats, publishers may use these flags in different letter cases: all in lowercase (e.g., "retracted: "), title case (e.g., "Retracted"), and all in uppercase (e.g., "RETRACTED"). Elango et al. reported that almost three-quarters of 239 (178/239) retraction notices included either "RETRACTION NOTICE TO" or "RETRACTED ARTICLE" (Elango et al., 2019).

Table 1: Examples of flags likely to indicate that an item is a retracted paper (Salami & Schneider, 2024)

	Flags Likely to Indicate Retracted Papers
Title starts with:	• [retracted] • retracted • retracted article • retracted: • retracted manuscript: • retracted:
Whole title is:	• retracted: • retracted
Title ends with:	• retracted

Table 2: Examples of flags likely to indicate that an item is a retraction notice (Salami & Schneider, 2024)

	Flags Likely to Indicate Retraction Notices
Title starts with:	• [retraction] • retraction notice: • correction to: retracted article: • correction to: • correction: • corrigendum to • removed: • retraction notice for • retraction note to: • statement of retraction for • temporary removal:
Whole title is:	• partial retraction notice • removal statement • retraction note • statement of retractions
Title ends with:	• :retraction • retraction

Flag indication is the interpreted or inferred publication status (e.g., retracted paper, retraction notices) derived from the title flag appearing in the title. The title flag "RETRACTED" suggests that the publication is a retracted publication, whereas the title flag "RETRACTION NOTICE" suggests that the publication is a retraction notice.

Publisher indication is the publication's retraction status as reported on the publisher's website.

As shown in Figure 2, the inferred retraction status from the paper’s title flag may not agree with the publisher indication.

Figure 2: This retraction notice has the "Retracted" title flag, which often indicates a retracted paper. In this case, the flag indication and publisher indication do not agree.

3. Research Questions

RQ1: Do DOIs indexed as retracted in Crossref (as of April 2023) have a publisher indication of retracted on the publisher's website?

RQ2: Do the publisher indication and the flag indication agree regarding whether a publication is retracted?

RQ3: What is the publisher indication for DOIs no longer indexed as retracted (as of July 2024) in Crossref?

4. Data and Methods

For this work, we curated two datasets. Both were drawn from our prior work, which created a union list of 49,924 purportedly retracted publications indexed as retracted in one or more of Crossref, Retraction Watch, Scopus, and Web of Science, as of April 2023 (Schneider et al., 2023). The first dataset, which is expected to have low retraction indexing data quality, is comprised of all 9937 DOIs indexed as retracted in Crossref but not in the other three sources (Lee & Schneider, 2023). The second dataset, produced by Salami et al., (2024) is expected to identify non-retracted items: It is comprised of all 427 DOIs from Schneider et al. (2023) that were indexed as retracted in one or more of Crossref, Retraction Watch, Scopus, and Web of Science as of April 2023, but are no longer so as of July 2024.

4.1 Methods for RQ1 & RQ2

For RQ1 and RQ2, we selected a sample from Lee & Schneider (2023). Of Lee & Schneider's 9937 DOIs, 1753 that had title flags, and from these, we selected a sample of 925 DOIs. The sample was drawn from all 715 journals, with the selection criteria: one DOI per title flag, or two

when the multiple DOIs within same journal shared the same title flag. This sample comprises two groups of flag indications: 331 DOIs with title flags suggesting retracted papers (see Table 1) and 594 DOIs with title flags suggesting retraction notices (see Table 2*)*.

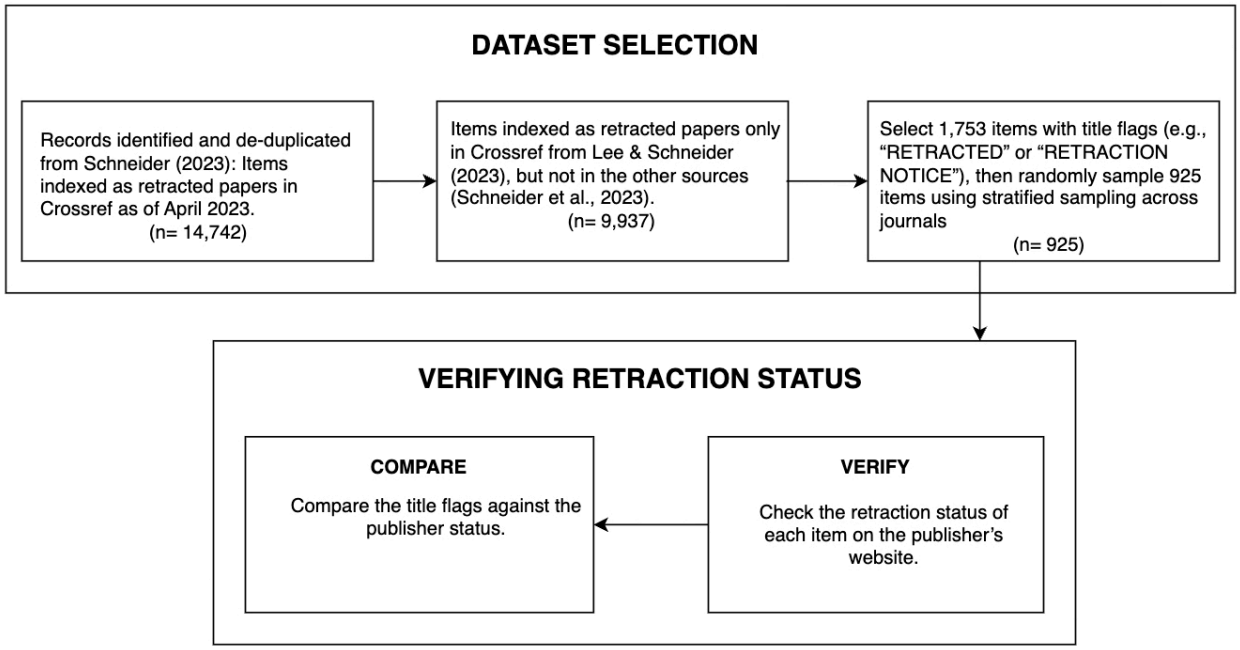
For RQ1, we manually checked the publisher indication for each DOI, classifying it as a retracted paper, retraction notice, or both retracted paper and retraction notice (see Si et al., 2024 for details).

For RQ2, we compared two different indications: the flag indication and the publisher indication to see whether they matched or were mismatched.

4.2 Methods for RQ3

To address RQ3, of the 427 DOIs from Salami et al., (2024), we selected all 208 from Crossref. First, we re-retrieved metadata for these DOIs via the Crossref API on August 27, 2024. Second, we manually checked the publisher indications for these 208 DOIs on their publishers' websites. Third, on July 8, 2025, we queried the Crossref API to check whether the title metadata had been updated, particularly whether the flag indication had changed or was no longer included in the title metadata.

Figure 3. Workflow for RQ1 & RQ2.



5. Results

5.1 RQ1: Do DOIs indexed as retracted in Crossref (as of April 2023) have a publisher indication of retracted on the publisher's website?

Table 3 summarizes the publisher indications. Of the sample of 925 DOIs indexed as retracted papers by Crossref, 71.46% were designated as retraction notices, while 12.86% were designated as retracted papers according to the publishers' websites.

Table 3: Publisher indication of sampled DOIs

Publisher indications	**Count (%)**
Retraction notice	661 (71.46%)
Both retracted paper and retraction notice	138 (14.92%)
Retracted paper	119 (12.86%)
Correction	3 (0.32%)
Unknown	2 (0.22%)
Not Available	2 (0.22%)

For DOIs whose publication indications are both retracted papers and retraction notices, almost all had title flags (98.55%, 136/138) that indicate retracted papers. These title flags are: "retracted article:" (99), "retracted:" (36), and "retracted article" (1). In only in two instances does the title flag indicate a retraction notice: "retracted notice:" (1) and "statement retraction" (1).

5.2 RQ2: Do the publisher indication and the flag indication agree regarding whether a publication is retracted?

Table 4 shows a comparison between the flag indications and publisher indications for 780 DOIs: 585 retraction notices and 195 retracted papers inferred based on their flag indications. For simplicity, we excluded the 138 items designated as both retracted papers and retraction notices DOIs from the analysis for RQ2.

Table 4: Comparison of publisher indication and flag indication

	Flag and publisher indications are the same	**Flag and publisher indications are different**
Retracted paper flag (N=195)	118 DOIs (60.51%)	77 DOIs (39.49%)
Retraction notice flag (N=585)	584 DOIs (99.83%)	1 DOI (0.17%)

Table 4 shows that when the title flag suggests a retraction notice, there is almost always (584/585, 99.83%) a match with the publisher indication. More than half of the title flags suggesting retracted papers (118/195, 60.51%) also match the publisher indication.

5.1 Retracted Paper Flags

We report on retracted paper flags that reflect the same retraction status (Table 5) and different retraction status (Table 6) as publishers.

The "retracted:" and "retracted article:" title flags are the most common variants (175/195, 89.74%) used to label retracted publications. However, their usage appears inconsistent and not reliable, as shown in Tables 5 and 6.

Table 5: Retracted paper flag indications in the sample that agree with the publisher indications.

Flag and publisher indications are the same (Retracted Paper Flag N=195)		
Flag	**# DOIs**	**Percentage (%)**
retracted:	71	36.41%
retracted article:	44	22.56%
retracted: corrigendum to	2	1.03%
retracted:	1	0.51%

Table 6: Retracted paper flag indications in the sample that are different from the publisher indications.

Flags and publisher indications are different (Retracted Paper Flag N=195)		
Flag (a bulleted list indicates multiple flags)	**# DOIs**	**Percentage (%)**
retracted:	40	20.51%
retracted article:	20	10.26%
retracted chapter:	6	3.08%
retracted, retracted: corrigendum to	3	1.54%
retracted: retraction notice to	2	1.03%
• [retracted] • retracted: erratum to • retracted manuscript:	1	0.51%

5.2 Retraction Notice Title Flags

Almost all the DOIs (99.83%, 584/585) with retraction notice flags matched the publisher indications as of July 2025, as shown in Table 7. The flags "retraction", "statement of retraction", and "retraction notice" were common in retraction notice titles. We observed 30 different title flag variants for retraction notices, including rare variants such as "partial retraction notice", "NOTICE", and misformatted HTML such as "retraction" (without the corresponding closing tag), as shown in Table 7.

Most retraction notice flag indications (90.26%, 528/585) were used on their own as the full title, omitting the original title of the retracted article. This practice is suboptimal, especially when full titles of retraction notices such as "CORRECTION" and "CORRIGENDUM" (2/250 cases, as shown in Table 7). This could lead to the incorrect interpretation of these retraction notices as correction notices.

Table 7: Retraction notice flag indications in the sample that match the publisher indications

Flags and publisher indications are the same (Retraction Notice Flag N=585)		
Flag (a bulleted list indicates multiple flags)	**# DOIs**	**Percentage (%)**
retraction	205	35.04%
statement of retraction	163	27.86%
retraction notice	99	16.92%
statement of removal	30	5.13%
editorial retraction	16	2.74%
retraction statement	12	2.05%
retracted article: statement of retraction	8	1.37%
• notice of retraction • retraction note	6	1.03%
• [retraction] • retraction • editorial removal • notice of redundant publication • retracted article: retraction • retractions	3	0.51%
• retraction notice: • retraction • correction • corrigendum • removal statement • statement of retractions	2	0.34%
• statement of retraction: • author's retraction • notice • partial retraction notice • publisher’s note • retraction: • retraction of articles • retraction statement: • revise and republish notice	1	0.17%

RQ3: What is the publisher indication for DOIs that are no longer indexed as retracted in Crossref as of July 2024?

Of the 208 DOIs that changed indexing in Crossref in April 2023 versus July 2024, none have been retracted according to the publishers. Interestingly, as shown in Table 8, 20 (9.6%) of these 208 DOIs are now normal, that is regular articles without any historical error notice or traces on the publishers' websites.

Table 8: Status of the 208 DOIs that are no longer indexed as retracted on the publisher website.

Publisher Indications	**Count (%)**
Retraction notice	176 (84.62%)
Normal article	20 (9.62%)
Correction notice	12 (5.77%)

19 of the 20 normal articles had their title flags updated, as shown in Table 9; many of these DOIs were for articles that were removed under "TEMPORARY REMOVAL" title flags, but were restored, and republished as normal articles (i.e., articles published without flags in their titles). Interestingly, of the 176 retraction notices, no changes had been made in their title metadata: 7 have no retraction notice flag in their titles, and 168 had variant retraction notice flags, except in one case: One DOI's title changed from "RETRACTED: Investigation of ductile fracture behavior of lap-welded joints with 460 MPa steel" to "RETRACTION NOTICE: Investigation of ductile fracture behavior of lap-welded joints with 460 MPa steel".

Table 9: Titles of 20 DOIs that Crossref no longer indexes as retracted are changed

Publisher Indications	**# DOIs**	**Changes in title**
Normal article	19	"removed" removed (2) "temporary removal" removed (16) "withdrawn" removed (1) no retraction flag in the previous title and present title (1)
Retraction notice	1	from "retracted" to "retraction notice"

6. Discussion

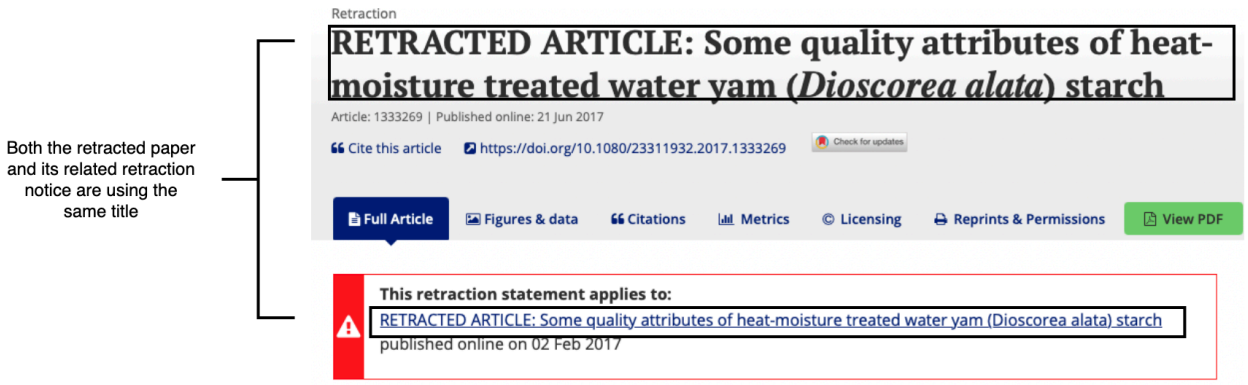
Our findings demonstrate that retraction metadata are dynamic rather than static. Between April 2023 and July 2024, Crossref removed retraction status from 208 DOIs, and our manual verification confirmed that these items were not retracted publications according to the publisher. The title flags for these 208 DOIs were subsequently updated or removed, which we deem data quality improvements

Overall, we found significant disagreements between Crossref and publisher indications regarding retraction status DOIs. According to publisher websites, 661 DOIs (71.46%) in our sample are retraction notices but indexed as retracted papers in Crossref. This discrepancy highlights issues in retraction metadata quality.

In addition, 138 DOIs (14.92% of the sample) have a publisher indication of both a retracted publication and a retraction notice, meaning that the retraction notice and retracted publication are associated with the same DOI on the publisher website. This means that their retraction notices were appended directly to the retracted (original) publication. This is bad practice, because they share the same DOI, even though Crossref (2025) recommends that retraction notices be registered as separate records with their own DOIs.

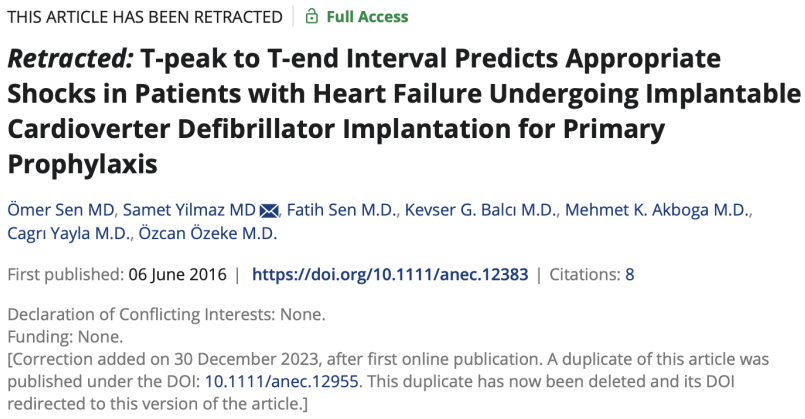
Retracted paper title flags do not reliably identify retracted papers. Among 195 DOIs that have retracted paper flag indications, 77 (39.49%) have flag indications that differ from the publisher indications. The most commonly used retracted paper flags, such as "retracted:" and "retracted article" (175/195, 89.74%), have the same publisher indications for only a subset of DOIs. Thus, these flags are not helpful to reliably infer a DOI's retraction status. One reason we observed for this unreliability is that some publishers assign the same title flag (despite different DOIs) to both the retracted paper and its retraction notice, as shown in Figure 4.

Figure 4: A retracted paper and its retraction notice both use the same title (DOI:10.1080/23311932.2017.1333269).



Retraction notice flags seem more reliable. The single discrepancy (Figure 5) between the retraction notice flag indications and publisher indications found in the July 2025 data was subsequently resolved. As of September 2025, all retraction notice flag indications agreed with their publisher indications. Thus, at this point, the retraction flags covered in the sample were all reliable. Given this reliability, retraction notice flags could be used for broader data quality improvements. Databases like Crossref could potentially verify the status of DOIs using retraction notice flags and update the indexing if those DOIs are mistakenly indexed as retracted papers.

Figure 5: Publisher information indicates a changed title and retraction notice


THIS ARTICLE HAS BEEN RETRACTED | Full Access

***Retracted:* T-peak to T-end Interval Predicts Appropriate Shocks in Patients with Heart Failure Undergoing Implantable Cardioverter Defibrillator Implantation for Primary Prophylaxis**

Ömer Sen MD, Samet Yilmaz MD, Fatih Sen M.D., Kevser G. Balcı M.D., Mehmet K. Akboga M.D., Cagrı Yayla M.D., Özcan Özeke M.D.

First published: 06 June 2016 | https://doi.org/10.1111/anec.12383 | Citations: 8

Declaration of Conflicting Interests: None.
Funding: None.
[Correction added on 30 December 2023, after first online publication. A duplicate of this article was published under the DOI: 10.1111/anec.12955. This duplicate has now been deleted and its DOI redirected to this version of the article.]

As shown in Figure 6, the "retraction" title flag has been used as a full item title without including the title of the retracted article. However, recent NISO guidelines recommend that "The retraction notice title should clearly identify the retracted publication" (NISO Communication of Retractions, Removals, and Expressions of Concern Working Group, 2024) while the 2025 COPE guideline also recommend that retraction notices should clearly identify the title and authors of the paper being retracted (COPE Council, 2025).

Figure 6: This retraction notice is simply titled "Retraction".


RETRACTION

Retraction

Page 1 | Published online: 28 Dec 2022

Cite this article https://doi.org/10.1080/10717544.2022.2157154 Check for updates

Full Article Figures & data Citations Metrics Licensing Reprints & Permissions

View PDF View EPUB Share

This retraction statement applies to:

RETRACTED ARTICLE: Potential of solid dispersions to enhance solubility, bioavailability, and therapeutic efficacy of poorly water-soluble drugs: newer formulation techniques, current marketed scenario and patents published online on 18 Nov 2020

Some flags were used ambiguously in the title. For example, we found that some titles start with "RETRACTED ARTICLE: Retraction" or "RETRACTED ARTICLE: Statement of Retraction". This is confusing! It is likely a retracted paper (based on the first flag, "RETRACTED ARTICLE")

AND likely a retraction notice (based on the second flags, "Retraction" and "Statement of Retraction", respectively). Further, retraction notice flags may vary within a journal. For example, *Applied Artificial Intelligence* has used both "retraction statement" and "statement of retraction". Such ambiguity complicates the use of title flags for automated classification in the bibliographic database system (see, for instance, Uppala et al., 2022) and may contribute to inconsistencies between flag indications and publisher indications.

Despite these ambiguities and inconsistencies in title flag usage, we also found some improvements in retraction metadata quality. According to publishers' websites, the 208 DOIs that had Crossref indexing changes (from retracted to not retracted) are indeed normal (non-retracted) papers. Therefore, data quality was improved by no longer indexing these DOIs as retracted papers. Notably, we found a case of temporary withdrawal whose title was changed, removing the "withdrawn" title flag (10.1016/j.gim.2022.09.015) from its title, so it appears as a normal article; it now has a correction notice (10.1016/j.gim.2023.100828). Researchers using metadata snapshots may obtain different results about retraction status depending on when the data are collected.

7. Limitations and Future Work

This study focuses only on Crossref. Future work could test flag reliability in other databases, such as by linking retracted publications and retraction notices to detect mistakes in retraction indexing (for a model, see Schmidt (2018)'s Web of Science analysis). While our work does not directly determine how inconsistencies in retraction-related metadata occur, future work could map data flows from publishers to aggregators to identify potential process improvements. Future work could also examine the association between flag location (beginning, end, full title) and its reliability in identifying retraction status. Assessing data quality changes longitudinally, over time, would also be beneficial.

8. Conclusions

Retraction metadata should be treated as dynamic rather than definitive, so periodic validation is necessary for high-stakes research and evaluation. Retraction notice title flags could be useful for assessing data quality, particularly to identify retraction notices mistakenly indexed as retracted publications. Some changes in Crossref's retraction indexing have improved data quality. When items are retracted, metadata updates should be timely, and NISO recommends that item titles change accordingly (NISO Communication of Retractions, Removals, and Expressions of Concern Working Group, 2024).

We highlight four main findings. First, we identified shortcomings in Crossref's retraction indexing practices: Many DOIs are indexed as retracted even though they correspond to retraction notices on publisher websites. Second, we showed that title flags are inconsistently used to distinguish between retracted publications and retraction notices: While title flags identifying retraction notices are largely reliable, those identifying retracted papers do not always match their publisher indications. Third, we showed that, within a bibliographic database system, retracting a publication is not definitive or permanent action, due to errors. Items misindexed as retracted may later be updated, and their associated title flags may change over time. Finally, we revealed a lag in communication and updates between publishers and Crossref in propagating changes to retraction classification and their title flag metadata.

Scientometrics researchers who study retractions using historical database snapshots and bibliographic databases (such as OpenAlex[3]) that ingest data from Crossref should verify the retraction status of publications. Although retracted papers and corresponding retraction notice should not share an identical DOI, in practice, be aware that some do. Large-scale studies can use auditing techniques: (1) cross-check title flags with retraction status and (2) identify retraction notices for all purportedly retracted publications. Scientometrics researchers should periodically reassess records whose retraction status has changed to prevent the propagation of outdated classifications.

Retrieving and reading the retraction notice is the best way to validate retraction status for small scale studies. Publishers' and databases' inconsistent practices can result in retracted publications being missed during literature searches or, conversely, in non-retracted items being incorrectly flagged as retracted. Such errors are most consequential for researchers conducting large-scale bibliometric analyses, systematic reviews, or evidence synthesis, librarians supporting research integrity and collection management, and developers of tools that rely on automated retraction metadata detection, as in the case of reference management tools.

Overall, our findings suggest that longitudinal monitoring of retraction metadata may be necessary to understand both the prevalence of indexing errors and the speed with which changes propagate across the scholarly communication ecosystem. We recommend that publishers and Crossref take shared responsibility for ensuring the accuracy and timeliness of retraction metadata.

Open science practices

Code is available in the infoqualitylab/tracking-data-quality-landscape-of-retracted-papers repository, archived at http://doi.org/10.5281/zenodo.20805350

We share our two datasets collected from the Crossref API:
Si, Luyang; Salami, Malik Oyewale; Schneider, Jodi (2025): Dataset on tracking the data quality landscape of retracted papers: Flag usage in titles and changes in DOI retraction status. University of Illinois Urbana-Champaign. https://doi.org/10.13012/B2IDB-2907908_V1

Si, Luyang; Salami, Malik Oyewale; Schneider, Jodi (2025): Dataset tracking DOIs marked as retracted in Crossref as of April 2023 but no longer marked in Crossref as retracted as of July 2024. University of Illinois Urbana-Champaign. https://doi.org/10.13012/B2IDB-5333456_V1

Acknowledgements

Thanks to Jou (Laura) Lee and Corinne McCumber for contributing to data and to Heng Zheng for an earlier manual analysis published in STI2023 that laid the groundwork for this analysis. We appreciate feedback from Corinne McCumber, Dexter Williams, Emily Wegrzyn, Hangyue Zhang, Jena Lui, Kiel Gilleade, Puranjani Das, Xiaochen Ma, Xiuping Shen, Yuanxi Fu and the Illinois Writers Workshop. We used Grammarly for editing support.

[3] https://help.openalex.org/hc/en-us/articles/24347019383191-Where-do-works-in-OpenAlex-come-from

Author contributions

Luyang Si: Data Curation, Formal analysis, Validation. Visualization Writing-original draft, Writing-review & Editing

Malik Oyewale Salami: Conceptualization, Data Curation, Formal analysis, Funding acquisition, Supervision, Validation, Visualization, Writing-original draft, Writing-review & Editing – lead

Jodi Schneider: Conceptualization, Funding acquisition, Supervision, Validation, Visualization, Writing-original draft, Writing-review & Editing

Competing interests

Malik Salami and Luyang Si declare no competing interests. Jodi Schneider is a member of the NISO Communication of Retractions, Removals, and Expressions of Concern Working Group, which was derived from her previous research as PI of Alfred P. Sloan Foundation G-2020-12623. Jodi Schneider declares non-financial associations with Crossref; COPE; International Association of Scientific, Technical and Medical Publishers; the National Information Standards Organization; and the Center for Scientific Integrity (parent organisation of Retraction Watch). The National Information Standards Organization was a subawardee on her Alfred P. Sloan Foundation grant G-2022-19409. She serves on the National Academies of Sciences, Engineering, and Medicine Consensus Study Committee, Corrections and Retractions: Upgrading the Scientific Record.

Funding information

This project was funded in part by Alfred P. Sloan Foundation G-2022-19409 Reducing the Inadvertent Spread of Retracted Science II: Research and Development towards the Communication of Retractions, Removals, and Expressions of Concern. JS was supported in part as the 2024-2025 Perrin Moorhead Grayson and Bruns Grayson Fellow, Harvard Radcliffe Institute for Advanced Study and conducted parts of this work while at the University of Illinois Urbana-Champaign and at the Harvard Radcliffe Institute for Advanced Study.

References

Bakker, C. J., & Riegelman, A. (2018). Retracted publications in mental health literature: Discovery across bibliographic platforms. *Journal of Librarianship and Scholarly Communication*, *6*(1), eP2199. https://doi.org/10.7710/2162-3309.2199

Bakker, C. J., Reardon, E. E., Brown, S. J., Theis-Mahon, N., Schroter, S., Bouter, L., & Zeegers, M. P. (2024). Identification of retracted publications and completeness of retraction notices in public health. *Journal of Clinical Epidemiology*, *173*. https://doi.org/10.1016/j.jclinepi.2024.111427

COPE Council. (2025). *COPE Guidelines for retracting articles [2025]*. https://doi.org/10.24318/cope.2019.1.4

Crossref. (2025, April 17). *Version control, corrections, and retractions* [Website]. Crossref. https://www.crossref.org/documentation/principles-practices/best-practices/versioning/

De Cassai, A., Geraldini, F., De Pinto, S., Carbonari, I., Cascella, M., Boscolo, A., Sella, N., Monteleone, F., Cavaliere, F., Munari, M., Garofalo, E., & Navalesi, P. (2022). Inappropriate citation of retracted articles in anesthesiology and intensive care medicine publications. *Anesthesiology*, *137*(3), 341. https://doi.org/10.1097/ALN.0000000000004302

Elango, B., Kozak, M., & Rajendran, P. (2019). Analysis of retractions in Indian science. *Scientometrics*, *119*(2), 1081–1094. https://doi.org/10.1007/s11192-019-03079-y

Lee, J., & Schneider, J. (2023). *Crossref data for Assessing the agreement in retraction indexing across 4 multidisciplinary sources: Crossref, Retraction Watch, Scopus, and Web of Science*. https://doi.org/10.13012/B2IDB-9099305_V1

Minetto, S., Pisaturo, D., Cermisoni, G. C., Vanni, V. S., Pagliardini, L., Papaleo, E., Berghella, V., Mol, B. W., & Alteri, A. (2024). Are you aware of your citations? A cross-sectional survey on improper citations of retracted articles in assisted reproduction. *Reproductive BioMedicine Online*, *49*(5), 104366. https://doi.org/10.1016/j.rbmo.2024.104366

NISO Communication of Retractions, Removals, and Expressions of Concern Working Group. (2024). *Communication of Retractions, Removals, and Expressions of Concern (CREC)* (NISO RP-45-2024). NISO. https://doi.org/10.3789/niso-rp-45-2024

Salami, M. O., McCumber, C., & Schneider, J. (2024, November 18). *Reassessment of the agreement in retraction indexing across 4 multidisciplinary sources: Crossref, Retraction Watch, Scopus, and Web of Science*. Zenodo. ASIS&T MET-STI 2024: Workshop on Informetric, Scientometric and Scientific and Technical Information Research. https://doi.org/10.5281/zenodo.14004526

Salami, M. O., & Schneider, J. (2024). *PROTOCOL: Distinguishing Retracted Publications from Retraction Notices in Crossref data [version 2.1]*. https://hdl.handle.net/2142/125231

Schmidt, M. (2018). An analysis of the validity of retraction annotation in PubMed and the Web of Science. *Journal of the Association for Information Science and Technology*, *69*(2), 318–328. https://doi.org/10.1002/asi.23913

Schneider, J., Lee, J., Zheng, H., & Salami, M. O. (2023, September 28). *Assessing the agreement in retraction indexing across 4 multidisciplinary sources: Crossref, Retraction Watch, Scopus, and Web of Science*. 27th International Conference on Science, Technology and Innovation Indicators (STI 2023). https://doi.org/10.55835/6441e5cae04dbe5586d06a5f

Si, L., Salami, M. O., & Schneider, J. (2024). Distinguishing retracted publications from retraction notices in Crossref data [Poster]. Poster presented November 6, 2024 at the University of Illinois at Urbana-Champaign iSchool Research Showcase https://hdl.handle.net/2142/125134

Suelzer, E. M., Deal, J., Hanus, K., Ruggeri, B. E., & Witkowski, E. (2021). Challenges in identifying the retracted status of an article. *JAMA Network Open*, *4*(6), e2115648. https://doi.org/10.1001/jamanetworkopen.2021.15648

Uppala, A., Rosati, D., Nicholson, J. M., Mordaunt, M., Grabitz, P., & Rife, S. C. (2022). *Title detection: A novel approach to automatically finding retractions and other editorial notices in the scholarly literature* (arXiv:2210.09553). arXiv. https://doi.org/10.48550/arXiv.2210.09553

Van Noorden, R. (2023). More than 10,000 research papers were retracted in 2023—A new record. *Nature*, *624*(7992), 479–481. https://doi.org/10.1038/d41586-023-03974-8

Wager, E., Barbour, V., Yentis, S., & Kleinert, S. (2009). Retractions: Guidance from the Committee on Publication Ethics (COPE). *Croatian Medical Journal*, *50*(6), 532–535. https://doi.org/10.3325/cmj.2009.50.532

Yan, J., MacDonald, A., Baisi, L.-P., Evaniew, N., Bhandari, M., & Ghert, M. (2016). Retractions in orthopaedic research: A systematic review. *Bone & Joint Research*, *5*(6), 263–268. https://doi.org/10.1302/2046-3758.56.BJR-2016-0047