

# Retraction Watch and Crossref: Collaborating to Improve the Assessment of Scholarly Outputs

Rachael Lammey and Ivan Oransky

## Introduction

Research integrity and issues of trust around scholarly publishing and the peer review process are under renewed scrutiny as editors deal with issues like plagiarism, artificial intelligence-generated papers, duplicate submissions, paraphrasing, image manipulation and duplication, fabricated data, and the challenge of paper mills that do all of these things.

While developing and improving screening techniques to identify these issues can struggle to keep up, it becomes increasingly important to update the scholarly record if issues are identified. This is a key piece of what journals and publishers are committing to do when they host and steward the publication of research outputs.

The information on updates to the scholarly record such as corrections, retractions, or expressions of concern should be communicated quickly, clearly, openly, and consistently by the publisher, and they should also be made available in a machine-readable format so that downstream tools and services like abstracting and indexing services can reliably identify and share the current status of a work. This currently is not happening in a comprehensive way, which means that errors and updates can go undetected as research outputs are used and reused, and we risk the proliferation of misleading or incorrect information.<sup>1</sup> As a community, we need to take steps to correct this.

## Crossref and Retraction Watch

In September 2023, Crossref and Retraction Watch announced that Crossref has acquired the Retraction Watch

data and opened it to the scientific community,<sup>2</sup> and the two organizations will combine and publicly distribute data about tens of thousands of retracted research papers and grow the service together. This agreement fast-tracks the availability of an open, comprehensive, and accurate body of retractions and information on retractions for anyone to use.

The Retraction Watch data has initially been made available by Crossref in a .csv file format, as well as in the Crossref Labs API. In the future, however, the data will be integrated into Crossref's "main" REST API alongside metadata provided by Crossref members. The information will be updated on an ongoing basis as Retraction Watch continues to identify and share data on retractions and other important updates to content.

As the announcement explains, "the Center for Scientific Integrity (the organization behind the Retraction Watch blog and database) and the Retraction Watch blog will remain separate from Crossref and will continue their journalistic work investigating retractions and related issues; the agreement with Crossref is confined to the database only and Crossref itself remains a neutral facilitator in efforts to assess the quality of scientific works."<sup>3</sup>

## Why Is This Needed?

As of September 2023, Crossref had just under 14,000 retractions in the metadata registered by Crossref members. This is part of Crossref's Crossmark service,<sup>4</sup> which supports the collection of standard information on retractions and is accompanied by a button that publishers can place on their websites. Readers can click on the Crossmark button to see the current status of a work, based on the Crossref metadata. All Crossref metadata, including this information on retractions, is made openly available via the Crossref REST API.

Crossref has encouraged publishers to register this metadata by removing the Crossmark-specific fee in 2020 and by adding it as one of 12 key metadata elements that members can see on their Participation Reports.<sup>5</sup>

In comparison, in September 2023, the Retraction Watch database contained records for 43,000 retractions. This shows that retraction information is missing from publisher

Rachael Lammey (<https://orcid.org/0000-0001-5800-1434>) is Director of Product, Crossref. Ivan Oransky (<https://orcid.org/0000-0002-0746-9288>) is Co-founder, Center for Scientific Integrity.

*Opinions expressed are those of the authors and do not necessarily reflect the opinions or policies of the Council of Science Editors or the Editorial Board of Science Editor.*

<https://doi.org/10.36591/SE-D-4701-01>

## CONTINUED

metadata or provided in inconsistent ways. This makes it difficult for downstream services to use the Crossref data as it is markedly incomplete. It is important that a reader using these services does not assume that a paper has not been retracted because the data just is not there. As a community, pulling these two sources of information on retractions closer together and making them open provides one solution.

### What Can Be Done Because of This Transition?

This transition will enable many and varied benefits when assessing scholarly outputs.

Having an open source of comprehensive retraction information removes barriers for the community to do research about retractions. The Reducing the Inadvertent Spread of Retracted Science: Shaping a Research and Implementation Agenda (RISRS), led by Jodi Schneider at the University of Illinois at Urbana Champaign brought stakeholders together to look at the inadvertent spread of retracted research and make recommendations to improve this.<sup>6</sup> Opening the Retraction Watch data supports the work of this research team, which previously had to rely on closed or incomplete sources of retraction data. Since September, we have also seen institutions like the Hong Kong University of Science and Technology publish their own analysis of the full dataset.<sup>7</sup>

It also supports the integration of retraction information into more tools and services. Retraction Watch data was previously available to license for organizations to use in their products. Users may have seen Retraction Watch information in bibliographic tools like Zotero or in search services like Clarivate's Web of Science. By providing the Retraction Watch data openly and via an open license, it is easier for the data to be used in other downstream tools and services. Since making the database open, we have already seen an integration in the System Pro search service,<sup>8</sup> interest from GetFTR<sup>9</sup> in integrating retractions, and expect many more examples to follow. Industry groups are also exploring workflows that can push notifications to readers, authors, repositories, archives, and other stakeholders to actively let the community know if a paper has been retracted.

Perhaps most importantly, this agreement supports the sustainability of the important work that the team at Retraction Watch do and gives the team breathing room and the potential for growth as Ivan explains on the Retraction Watch blog.<sup>10</sup>

### What Still Has To Be Done?

Publishers remain a key, authoritative source of information on retractions and other important updates to content.

Both Crossref and the Center for Scientific Integrity see the Retraction Watch data as a complement to publisher-provided data on retractions. Better information on retractions from publishers supports all of the use cases listed in the previous section.

Publishers can check the Crossmark section of their Crossref Participation Report<sup>11</sup> to see if they are providing information on retractions and other updates to Crossref. If not, there is support documentation<sup>12</sup> that explains the additional metadata and information they can provide in order to do so, which they can share with their production teams or service providers.

We expect and hope that publisher provision of retraction information will accelerate as publishers implement more consistent and comprehensive retraction publication processes in line with the recommended practice from the NISO CREC (Communication of Retractions, Removals, and Expressions of Concern) Working Group,<sup>13</sup> which was released for public comment in October–December 2023 and will see a final version published in early 2024. There is still work to do by the whole community to better serve any reader wanting to know if a piece of content is current, but this is another important step toward that goal.

### References and Links

1. Hsiao T-K, Schneider J. Continued use of retracted papers: temporal trends in citations and (lack of) awareness of retractions shown in citation contexts in biomedicine. *Quant Sci Stud.* 2022;2:1144-1169. [https://doi.org/10.1162/qss\\_a\\_00155](https://doi.org/10.1162/qss_a_00155).
2. <https://www.crossref.org/blog/news-crossref-and-retraction-watch/>
3. Hendricks G, Lammey R. Crossref acquires Retraction Watch data and opens it for the scientific community. *Crossref Blog.* September 12, 2023. <https://doi.org/10.13003/c23rw1d9>.
4. <https://www.crossref.org/services/crossmark/>
5. <https://www.crossref.org/members/prepare/>
6. Schneider J, Woods ND, Proescholdt R; the RISRS Team. Reducing the inadvertent spread of retracted science: recommendations from the RISRS report. *Res Integr Peer Rev.* 2022;7:6. <https://doi.org/10.1186/s41073-022-00125-x>.
7. <https://library.hkust.edu.hk/sc/crossref-retraction-watch/>
8. <https://system-270d2a-c8ed176a779c27104e0350735.webflow.io/blog/introducing-retraction-watch-in-system-pro>
9. <https://www.getfulltextresearch.com/>
10. <https://retractionwatch.com/2023/09/12/the-retraction-watch-database-becomes-completely-open-and-rw-becomes-far-more-sustainable/>
11. <https://www.crossref.org/members/prepare/>
12. <https://www.crossref.org/documentation/crossmark/participating-in-crossmark/>
13. <https://www.niso.org/standards-committees/crec>