https://doi.org/10.51408/issi2025_097

A Dashboard to Visualize Retraction Statistics

Ayush Tripathi[1], Achal Agrawal[2], Moumita Koley[3]

[1] *yush.pbh@gmail.com*
Independent Researcher (India)

[2] *founder@irw.co.in*
India Research Watch (India)

[3] *moumitakoley@iisc.ac.in*
DST-Centre for Policy Research, Indian Institute of Science (India)

Abstract

Retraction Statistics are an important signal into studying research misconduct. We have created a dashboard to help visualize country-wise retraction statistics using the data from Retraction Watch Database. The dashboard helps view retraction rates of various countries over the years. The reasons for retractions are classified into various classes as described in a previously developed taxonomy of retractions. This tool can help journalists, policymakers as well as librarians to analyze retraction statistics. We plan to add more features like Institute-wise and author-wise analysis for every country. Institute-wise statistics can also be useful for ranking purposes. The dashboard can be accessed at https://retraction-dashboard.netlify.app/

Introduction

In 2023, there were more than 10,000 retraction notices, an all-time high (Van Noorden 2023). By some estimates, about 60% of those retractions are due to some form of research misconduct (Campos-Varela 2019). Thus, it is important to keep a close watch on the retraction statistics as they give us important clues about when and where research misconduct might be increasing to be able to take corrective actions.

A recent analysis of country-wise retraction rates found that Ethiopia had the highest retraction notice rate in the last 3 years (2022-2024) among the countries with at least 100 retractions in that time period (Agrawal 2025). This was the first time Ethiopia has been flagged in such a study and it is only possible when one monitors the statistics in permanent manner.

Many studies have reported on country-wise retraction statistics and drawn insights from them. Sharma (2024) studied retractions from past 2 decades in India. Shi (2023) did a regional analysis of retractions from China. It is clear that studying retractions can provide a lot of clues to the nature and location of misconduct.

Retractions are extremely tough to obtain, requiring 18 months on an average. For each paper that is retracted, there are many more that should be retracted. Heathers (2024) estimated that 1 in 7 science articles are fake or falsified. While a correct estimate is tough to obtain, there is consensus that retractions represent a very small fragment of misconduct. It is all the more reason why one must pay more attention to retraction statistics as they are an important signal.

With this dashboard, we provide updated retraction statistics for policymakers, journalists and librarians to analyze country-wise data to gain insights. In future, we plan to include institute-wise and author-wise statistics. Author disambiguation is done well in the retraction watch database and is fairly accurate. Institute disambiguation however is a tough problem as various versions of the Institute names are recorded in the database. We hope to solve this issue with the use of an external database.

Data and Methodology

We principally use two sources of data for the statistics displayed on the dashboard. For retractions, we use Retraction Watch Database (2018) which has the most number of retractions indexed. Crossref recently acquired the Retraction Watch Database and has made it open, enabling the creation of this dashboard. For country and year-wise number of publications we rely on SCImago (n.d.). While Retraction Watch Database also indexes articles which are not indexed in the Scopus database, SCImago includes only Scopus indexed publications. Thus, this is not fully accurate while calculating the retraction rate, but it does help a comparative analysis as the same method is applied uniformly to all countries.

To help better understand the reasons for misconduct, we classify all the reasons into various categories based on the Retraction Taxonomy developed by McIntosh (2024). Every retraction could have multiple reasons for retraction. The classification is done based on the priority of the reasons. There are 5 categories in the taxonomy in the order of priority:

Alterations: This category pertains to Data, Methods and Results. This includes concerning reasons like plagiarism, manipulation, falsification, duplication etc. This category is shown as red in the dashboard.

Author Integrity: This contains other form of misconduct like false peer review, ethical violations, lack of approvals, lack of ethics, conflict of interest etc. This category is shown as yellow in the dashboard.

Research: Sometimes, research could be retracted due to errors in the papers. These errors could be honest mistakes. It also contains reasons which make the research unreliable. This category is shown as blue in the dashboard.

System: This includes myriads of reasons pertaining to some issue at the system level like legal issues, miscommunication, objections or third-party violations. This category is shown as black in the dashboard.

Supplemental: This includes reasons like when papers are withdrawn by authors or if some investigations are initiated. These are fairly harmless bureaucratic reasons and is shown as grey in the dashboard.

The retraction dates used in the dashboard are the dates of original papers as is common practice while defining retraction rates. Another possibility is to use retraction notice rates, as done in Agrawal (2025), where dates of retraction notices are considered. They help provide a more recent signal of misconduct.

Features

For the design of the dashboard, we took inspiration from COKI Open Access Dashboard (Diprose 2023). We have three different types of pages for visualising the data: Main Dashboard, Country Page, and Comparison page.

Main Dashboard

Main dashboard contains a table with country-wise aggregate statistics of retractions under different categories. It also shows the retraction rate as well as the trend of retractions over the years in a compact form. The table can be sorted based on any column. Fig. 1 shows a screenshot of the first page of the dashboard.

Figure 1. Main Dashboard showing aggregate country-wise retraction statistics.

Country Page

Country page contains more detailed information about every country. It presents in-depth statistics of the evolution of the retraction rate, year-wise breakdown of different categories of retractions, as well as the countries collaborating in the papers which were retracted.

Fig. 2 shows the change in retraction rate for China over the years. It is interesting to see that there are two periods when there are sudden jumps in retraction rates. We can explore these jumps in Fig. 3 which shows the categories that the retractions in different years belong to. We see that in the period 2010-2011, many of the retractions are marked supplemental. These retractions are less worrying as they are mainly due to bureaucratic reasons. However, in the later period 2021-2022, there are more of the type Alterations and Researcher Integrity. These are worrying signs

for Chinese research. Chinese government has recently announced extensive investigations of the retractions and promised action against those found guilty of misconduct.

We can also see the countries collaborating in the papers which were retracted in Fig. 4. It can help understand the networks between different countries. Anomalous collaborations can provide connections between researchers of the countries to be investigated.

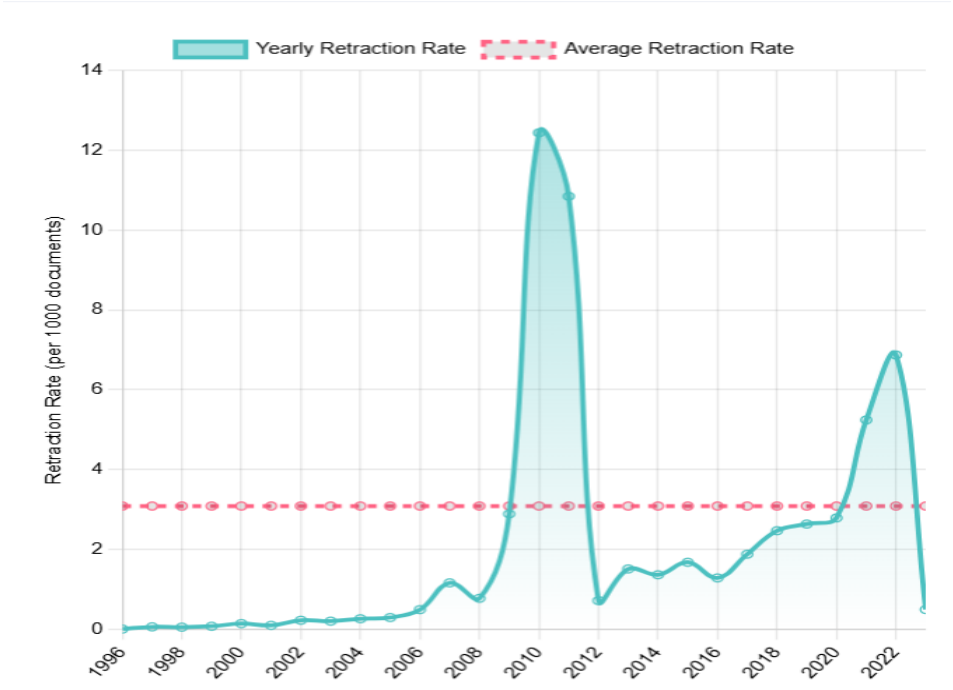

Figure 2. Retraction Rate over the years for China. We see two big jumps, once in 2011-2012 and other in 2021-2022.

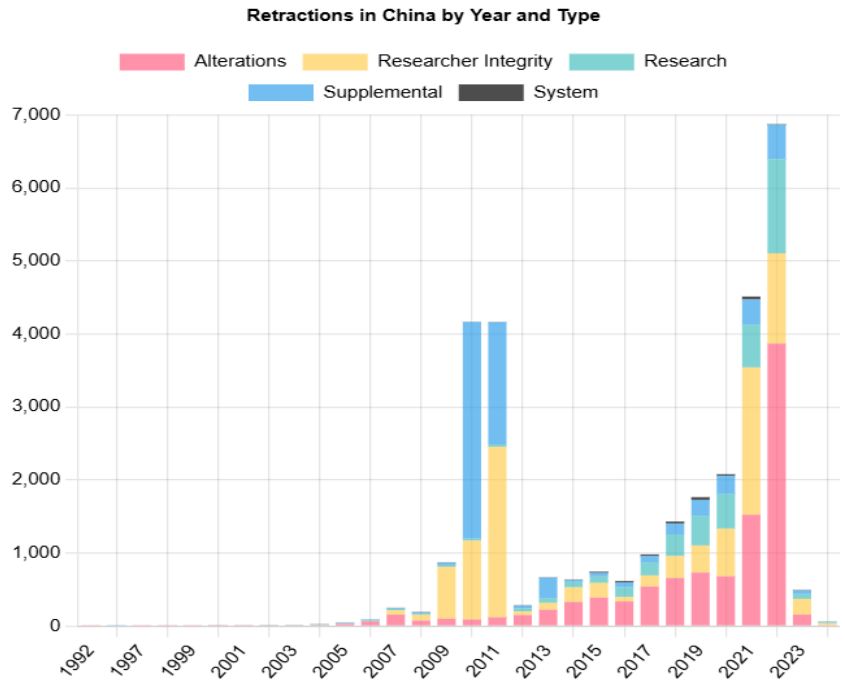

Figure 3. Retractions in China by year and category. First jump (2010-2011) is mostly due to supplemental reasons whereas the second jump (2021-2022) is more due to alterations.

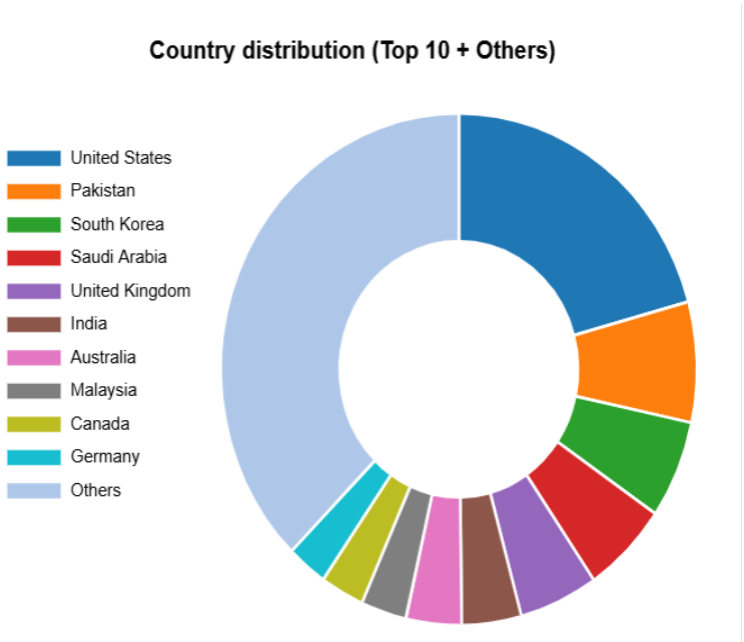

Figure 4. Countries collaborating with China in the papers which were retracted.

Comparison Page

In this page, one can choose various countries to compare with each other on a single graph. We plot the retraction rates of the chosen countries from 1996-2023. In Fig 5. we can see that India and China have increased their retraction rates greatly as compared to USA, UK and Japan.

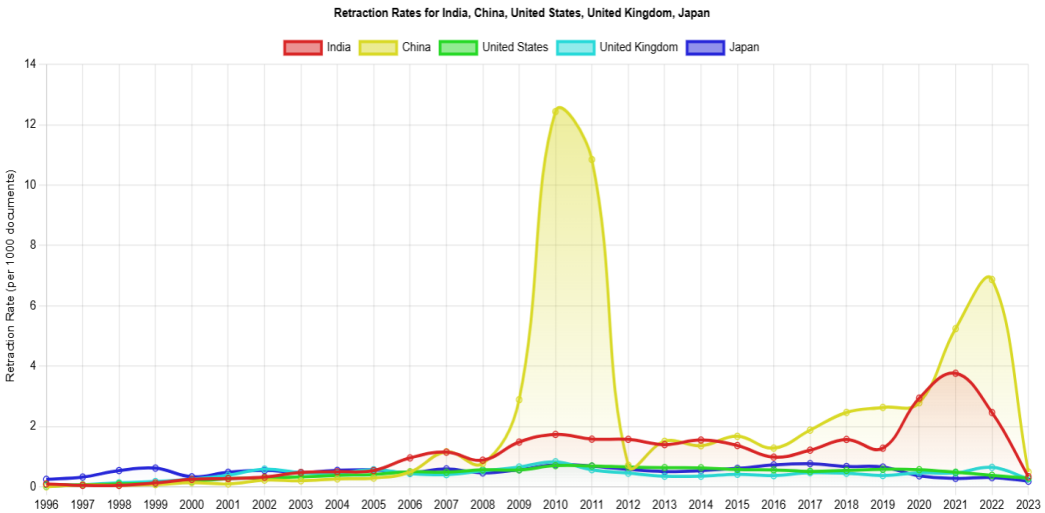

Figure 5. Comparison of retraction rates of various countries. We see China and India have an increased retraction rate lately.

Future Work

We are working to add many more features to the dashboard. Most new planned features are for the *Country Page*. We will add domain-wise, institute-wise and author-wise retraction data for every country. We also plan to provide Retraction Notice Rates, based on dates of retraction notices as they provide a more recent picture.

Additionally, we are also creating a notification system for universities to get alerted as soon as there is a retraction of any paper. Universities can update our system on various actions taken like investigation initiated and the decisions post the investigations. This is to help improve accountability of the universities to take retractions seriously and take appropriate actions.

Acknowledgments

We would like to thank Open Research Funders Group (ORFG) for providing the seed grant to develop this dashboard. Our project PostPub which builds upon this dashboard has been awarded the Catalyst Grant by Digital Science. The conference registration fees and travel have been funded by Digital Science.

References

Van Noorden, R. (2023). More than 10,000 research papers were retracted in 2023—a new record. *Nature*, *624*(7992), 479-481.Borgman, C.L. (Ed.). (1990). *Scholarly Communication and Bibliometrics*. London: Sage.

Campos-Varela, I., & Ruano-Raviña, A. (2019). Misconduct as the main cause for retraction. A descriptive study of retracted publications and their authors. *Gaceta sanitaria*, *33*, 356-360.

Agrawal, A. (2025). Country-wise Retraction Analysis from 2022-2024. Increased Publishing Leading to Higher Retraction Rates. Zenodo. https://doi.org/10.5281/zenodo.14634373

Sharma, K. (2024). Two decades of scientific misconduct in India: Retraction reasons and journal quality among inter-country and intra-country institutional collaboration. arXiv. https://arxiv.org/abs/2404.15306

Shi, L., Zhang, X., Ma, X., Sun, X., Li, J., & He, S. (2024). Mapping retracted articles and exploring regional differences in China, 2012–2023. *PloS one*, *19*(12), e0314622.

Heathers, J. (2024, September 24). HOW MUCH SCIENCE IS FAKE? https://doi.org/10.17605/OSF.IO/5RF2M

The Retraction Watch Database [Internet]. New York: The Center for Scientific Integrity. 2018. ISSN: 2692-4579. [Cited 9 Jan 2025]. Available from: http://retractiondatabase.org/.

SCImago, (n.d.). SJR — SCImago Journal & Country Rank [Portal]. Retrieved *9 Jan 2025*, from http://www.scimagojr.com

McIntosh, Leslie D.; Hudson Vitale, Cynthia (2024). Taxonomy of Retraction Reasons. figshare. Collection. (https://doi.org/10.6084/m9.figshare.c.7252732.v2)

Diprose, J., Hosking, R., Rigoni, R., Roelofs, A., Chien, T., Napier, K., Wilson, K., Huang, C., Handcock, R., Montgomery, L., & Neylon, C. (2023). A User-Friendly Dashboard for Tracking Global Open Access Performance. The Journal of Electronic Publishing 26(1). DOI: https://doi.org/10.3998/jep.3398