

What does it mean to correct the scientific record? A case study of the JACS (2000-2023)

Frédérique Bordignon

▶ To cite this version:

Frédérique Bordignon. What does it mean to correct the scientific record? A case study of the JACS (2000-2023). 27. International Conference on Science, Technology and Innovation Indicators (STI 2023), ENID, Sep 2023, Leiden, Netherlands. 10.55835/643e9f61c0a5d2717e40e59c. hal-04209423

HAL Id: hal-04209423 https://enpc.hal.science/hal-04209423

Submitted on 17 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



What does it mean to correct the scientific record? A case study of the JACS (2000-2023)

Frédérique Bordignon*1, 2

*frederique.bordignon@enpc.fr 0000-0002-4918-9137

¹ Ecole des Ponts, 77455 Marne-la-Vallée, France

Abstract

This paper examines how the *Journal of the American Chemical Society (JACS)* displays notices of correction and retraction, and how their status is reflected across various venues. With a corpus of 1083 editorial notices (2000-2023), we first show that even on the JACS website, the original source, there are mistakes and inaccuracies. Additionally, our study demonstrates some improvements in certain contexts in comparison to earlier studies, as well as significant variations between platforms (bibliographic databases and open access archives). It also reveals that the same types of issues still remain, including the lack of accurate information close to the updated publications, and the lack of a two-way link between notices and original publications. This preliminary research seeks to provide an overview of what constitutes the scientific record and what it means to correct it, in order to avoid the spread of unsubstantiated claims by ill-informed readers.

1. Introduction

In this paper, we consider the scientific record, especially the preservation of its integrity through editorial actions, i.e., retractions or publication of correction notices¹.

When errors or misconducts have been detected, it is essential to undertake the "cleaning of the published body of evidence" (Bouter, 2023) to limit the consequences that compromised or invalid results could have within the scientific community (e.g., with false leads, or funds dedicated to research doomed to fail), but also beyond with harmful usage in society (when used by policy makers, for instance).

Publishers are often reluctant to make such corrections for fear of compromising the reputation of their journal, but when they do, to what extent do they do it properly? How is the information disseminated across the various venues and databases?

This is what we propose to address in this paper, by presenting the results of a case study on a single journal, namely the *Journal of the American Chemical Society*.

2. Background

Hinchliffe (2022), when introducing the version of record as "a central organizing concept in scholarly publishing" reminds us of the NISO² definition: the version of record is "a fixed version of a journal article that has been made available by any organization that acts as a publisher by formally and exclusively declaring the article "published"" (NISO, 2008). We can observe the performative action of the publisher which, by publishing an article, contributes to the scientific record. Consequently, the publication of a correction or the retraction of a paper

² LISIS, INRAE, Univ Gustave Eiffel, CNRS, 77455 Marne-la-Vallée, France

¹ We do not include expressions of concern. Furthermore, we do not consider the difference between erratum and corrigendum, and between retraction and withdrawal, as the JACS does not make this difference.

² The American National Information Standards Organization.

constitutes an update of this record, which Dougherty (2019) describes as a "disruptive intervention", as opposed to the internal corrections, arguably more subtle, being part of the normal process of scientific progress, and referring to the ideal self-correcting capacity of science. While in both cases we refer to correcting the scientific record, the focus in our study is on updates performed by publishers (initiated by authors or not). It is critical that these updates should be easily findable, clear and prominently displayed.

In 1987, the Director of the National Library of Medicine in charge of the Medline bibliographic database already called for better indexing of retractions: "the general reader of the published scientific literature must be able to learn that an article he or she has read has subsequently been retracted" (Lindberg, 1987). At the same time, Garfield (1991) pointed out the crucial role of retraction and correction notices in scientific communication by presenting them as "important devices to ensure that science progresses on firm ground".

There is a long-standing consensus on the importance of editorial notices for the successful development of science and the imperative that they be easily accessible, visible and properly linked to the version they are updating. They have already been the subject of previous studies and we know they can be used to identify the different types of errors, ranging from typographic errors to invalid conclusions (Addelston & Goldsmith, 1966; Hubbard, 2010; Kiang, 1995; Sabine, 1985), and the reasons for correction/retraction (Casadevall et al., 2014; Coudert, 2019). They are also known to be of very different kinds depending on the journals and publishers (Jones et al., 2003; Teixeira da Silva & Vuong, 2022). They are known to resort to a euphemistic style (Hu & Xu, 2020). Finally, the link between the notice and the original paper it is intended to update is often not made properly or is missing (Jones et al., 2003; Poworoznek, 2003; Teixeira da Silva & Dobránszki, 2017).

Our analysis takes place in the context of the NanoBubbles project³, a European Research Council (ERC) Synergy project that aims to better understand the mechanisms of correction of science, focusing on bionanoscience. Considering the magnitude of the task of delineating the perimeter of the journals in the field of "nanos" and building a reliable corpus of publications linked to their possible correction or retraction notices, we chose to start with a particular case study that would constitute a first material. For reasons of consistency with other works in progress in the project, and a study already carried out by a member of the project (Noel, 2020), we chose a leading journal in the field of chemistry, the *Journal of the American Chemical Society*, more often referred to as "the JACS". The JACS is published weekly and is devoted to the publication of fundamental research papers in all fields of chemistry (*The ACS Guide to Scholarly Communication*, 2020). This journal has already been included in prior studies investigating the characterisation and quantification of corrections, error types or reasons of retraction (Addelston & Goldsmith, 1966; Sabine, 1985) and more recently by Hubbard (2010) whose set of 220 corrections published by the JACS between 2000 and 2005 matches with those we have identified for the same period, although we used a different methodology.

3. Methods

In line with the idea that the scientific record is generated and updated by publishers, we use the JACS website as the primary source to identify corrections that are tagged as *Addition/Correction* or *Retraction*. Insofar as we seek to evaluate their dissemination across other venues, we do not start by querying a bibliographic database as Hubbard (2010) and Jones *et al.* (2003) did by querying the Web of Science and Medline respectively. Even with this approach, by looking for potential corrections and retractions right at the source, it is not

³ https://nanobubbles.hypotheses.org

impossible that some are completely invisible because they have not been dealt with individually, and are listed in a set of errata, as was the practice in the days of paper journals. However, this risk is mitigated by the fact that our investigation focuses on the last 2 decades (while the JACS has been founded in 1879) in order to have a view of the ongoing correction practices.

On 8 April 2023, we collected 1083 correction notices published by the JACS since 2000, i.e., 1061 Additions/Corrections, and 22 Retractions. They point to 1068 publications (53% of which are Articles, 45% are Communications, and less than 2% are other document types). The dataset is available for download (Bordignon, 2023).

We therefore carried out the following processes:

- We manually checked the quality of the information provided by the JACS at the level of the original publication webpage (checking for the existence of a statement informing about an editorial notice and linking to it, either by a banner and an active URL or by a written mention) and at the level of the correction notice (with the same type of active banner linking to the corrected paper, or by a written mention). We read the notices to check whether the type (Addition/Correction or Retraction) is consistent with the content. We labelled the reporting of the information as "Adequate" when the Addition/Correction or the Retraction are really so, and when there is a back-and-forth link between the notice and the publication (or even several links, if multiple corrections have been successively performed). If one of these elements is missing, or is false (e.g., a broken link, a wrong categorization), we label it as "Inadequate".
- With the DOIs of the notices, we queried the Web of Science, Scopus and PubMed (three bibliographic databases) to check to what extent the notices are included and under what category. We considered the data presentation as "Adequate" when the Addition/Correction or the Retraction were really so (even if the JACS terminology is not used, such as *Erratum* or *Correction*), when there is a back-and-forth link between the notice record and the publication record (or if the status is correctly indicated in the title (e.g.: "RETRACTION OF: ...")). If one of these elements is missing, or is false (e.g., a broken link, a wrong categorization), we labelled it as "Inadequate". Finally, we distinguished between notices that are not included in the database, and those that are not included in the database *and* whose associated publication is also missing.
- With the DOIs of the original papers⁴, we queried the Web of Science, Scopus, PubMed, OpenAlex⁵ and Crossref⁶ to check for the presence of the correction and/or retraction statement(s) at the level of the description of each paper. We did not include Dimensions in the study as errata are not part of the information they provide, only retractions are indicated. For each database, we labelled the items as "Adequate" or "Inadequate"; we also collected the information that the publication is or is not included in the database.
- With the Unpaywall⁷ API, we retrieved the locations (URLs) of possible open access deposits and we manually scanned the records in search of correction mentions.

⁴ Except for 2 articles whose DOIs are missing, we thus used the title.

⁵ An open access database of scholarly works metadata.

⁶ An agency delivering DOIs.

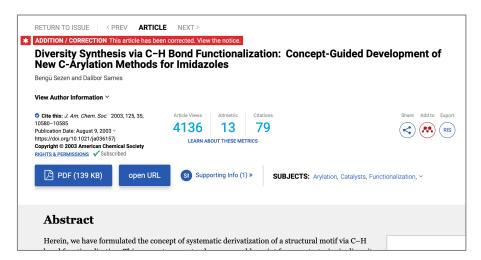
⁷ An index of scholarly research, including the link to an open access version of the full-text of any publication with a DOI.

4. Results

4.1. Presentation of the information on the JACS website

Of course, we agree that both errata and retractions are conceptually science correction processes. Nevertheless, they should be distinguished with different labels. Although the JACS differentiates between Additions/Corrections and Retractions, there are 18 corrections that are in fact retractions. This means that, strictly speaking, the correction of the scientific record was wrongly performed at the very source. These retractions therefore do not appear in the category of retractions. Yet their content leaves no room for doubt with clear statements such as "the authors retract...". This misclassification can deceive the user, who is certainly informed of a correction but not of a retraction as in Figure 1.

Figure 1. Screenshot of the JACS webpage for a retracted article (10.1021/ja036157j) with no mention of retraction



We fixed these problems in our corpus in order to "reconstruct" the scientific record as it should be presented by the JACS, and we are able to say that despite these shortcomings, 97.4% of the publications associated with a correction or a retraction notice are correctly displayed by the JACS (i.e., with a back-and-forth link to the notice and the correct mention of the status). For the part of the 2000-2005 corpus that aligns with Hubbard's (2010), we can even say that the situation has improved since his study in 2010, with a rate increasing from 96% to 99%, which means that corrective updates must have been made by the JACS over the past period.

With this reconstructed database, we are then able to verify the accuracy of the information provided by the bibliographic databases.

4.2. Dissemination of the information across bibliographic databases

The indexing of correction or retraction notices in bibliographic databases is of uneven quality depending on the database: in the Web of Science, coverage is good and there are few errors (0.8% over the whole period); in Scopus, coverage is good but there are more errors (2.6% over the whole period); and in PubMed, coverage went through ups and downs in the early 2000s but has improved significantly while the quality of reporting corrections has always been excellent over the full period.

Figure 2. Evaluation of the presentation of correction and retraction notices in the Web of Science (2000-2023)

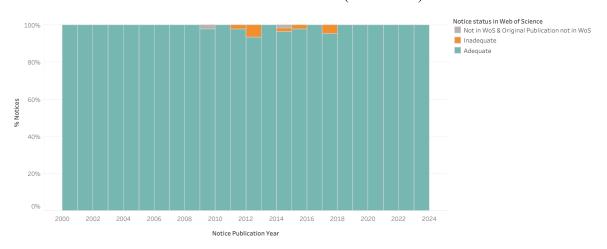


Figure 3. Evaluation of the presentation of correction and retraction notices in Scopus (2000-2023)

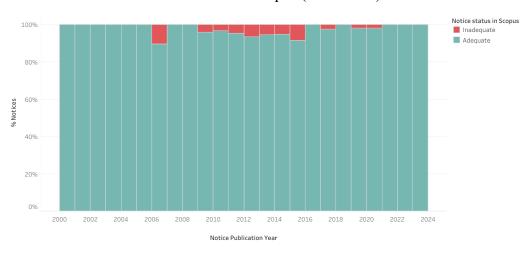
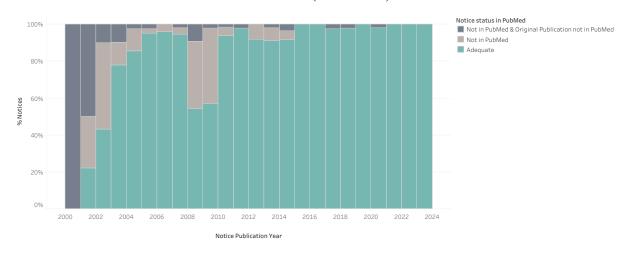


Figure 4. Evaluation of the presentation of correction and retraction notices in PubMed (2000-2023)



Now looking at the original publications, the first observation we make is that, surprisingly, what is not properly or prominently displayed on the JACS website is not necessarily

misreported in other venues. The most striking example is the following: out of the 18 publications miscategorized as corrections on the JACS website, 17 appear correctly in PubMed, and the same is true for 10 of them in the Web of Science (all belonging to the set of 17 rightly presented in PubMed). None are "corrected" in Scopus.

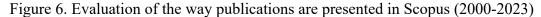
Over the whole corpus and period, the Web of Science is the most imprecise in that there is no possibility of being aware of the existence of a correction if one comes across the record of a publication. Only retractions are specifically flagged. Even if the correction exists elsewhere in the database, we consider the presentation of these publications to be inadequate, as the users are not sufficiently informed. They have to check all their results one by one by searching the titles in the database to eventually spot items categorized as corrections.

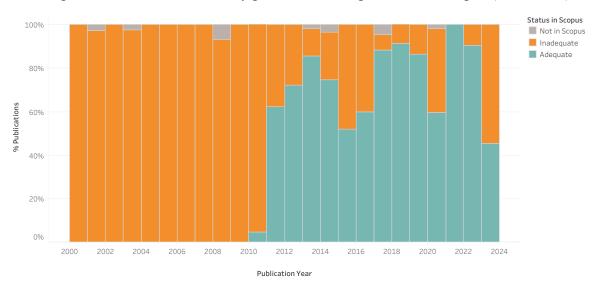
Status in Web of Science
Not in WoS
Inadequate
Adequate

40%
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
2024

Publication Year

Figure 5. Evaluation of the way publications are presented in the Web of Science (2000-2023)





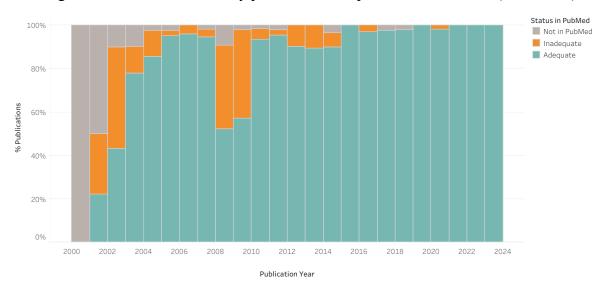


Figure 7. Evaluation of the way publications are presented in PubMed (2000-2023)

We were confident that Crossref could provide a solution to the shortcomings of these databases, especially since a metadata (*update-to*) dedicated to updates of the version of record exists and is used to feed the Crossmark service⁸. But on the one hand, Crossmark is not implemented by the publisher for the JACS. On the other hand, we have searched all the publications in Crossref⁹ (by looking for DOIs via the API), and none of them is linked to a correction information that should appear in the update field. The same is true for the OpenAlex database, whose data schema also includes a field (*is retracted*) dedicated to retractions.

4.3. Dissemination of the information across Open Access repositories

It is difficult to present an exact percentage distribution of results as some publications are available in open access on various platforms that we can identify with Unpaywall. But what can be said is that the major platforms, such as PubMed Central or EuropePMC, are the most accurate since they are probably directly fed by PubMed. However, they also contain errors (such as a notice in EuropePMC (DOI: 10.1021/ja108197s) which does not link to the correct document). We also came across the full-text of a retracted publication (DOI: 10.1021/ja201074e) in Figshare (a multidisciplinary repository), without any mention of retraction, and uploaded after the date of retraction.

As for the institutional repositories, they very rarely mention the existence of a correction. We have identified only 5 mentions of a correction. But this does not mean that the mention *should* really be there; indeed, it is possible that the deposited version deviates from the version of record and does not contain the error finally published in the final version...

These repositories are most probably maintained by librarians after the self-archiving of the documents by authors. It is a huge and unachievable task to track corrections and retractions, and report them manually at the level of each repository (since a connection to Crossref would not be reliable). Ironically, this is a replication of the problem librarians used to face in managing paper holdings by manually reporting corrections in libraries (Cooper, 1992).

⁸ "Crossmark is a set of relationship metadata and a button on article pages and PDFs that alerts readers to the currency and accuracy of the object they're looking at." (Hendricks et al., 2022)

⁹ Note: this study was performed before Crossref announced the acquisition of Retraction Watch data; see https://www.crossref.org/blog/news-crossref-and-retraction-watch/ (September 11th, 2023)

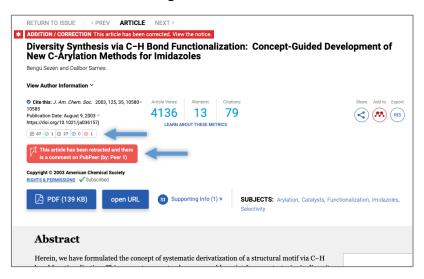
Dougherty sees this problem, which he termed the *Repository Problem*, as "arguably the most significant threat to keeping researchers up to date about any change in the status of a work" (Dougherty, 2019).

5. Discussion

The limited corpus we have built up and the focus on one journal obviously do not allow generalization to the entire body of scholarly literature. But it does have the merit of demonstrating the magnitude of the task when one aims to correct the scientific record, within a global infrastructure whose robustness of linkage is a strength when everything is going well, but which sometimes becomes a hindrance when it comes to "going backwards".

These preliminary results also challenge our beliefs about what is the scientific record, and deviate from the NISO definition. Insofar as the source cannot be trusted because it contains errors, we are forced to consider the scientific record as an aggregation of data from multiple providers, the validity of which is based on a bundle of information from legitimate sources. The different databases and information providers seem to feed (or even correct) each other. But since no source seems to be perfect, it is still up to the researchers to check the validity of information. This brings us back to Garfield's call: "Scientists ought to develop the habit of looking for corrections or retractions of works they cite in their publications" (Garfield, 1991). This habit is nowadays much easier to implement with the many technical connections that are developed between the components of scientific communication. For example, in addition to an active query in the bibliographic databases, the problematic case shown in Figure 8 is solved by the PubPeer¹⁰ and Scite¹¹ extensions at the browser level:

Figure 8. Screenshot of the JACS webpage for a retracted article (10.1021/ja036157j) with no mention of retraction but warnings from Scite and PubPeer browser extensions



In this context, the infrastructure as a whole, and in particular the technical connection established by and between the identifiers (e.g., DOIs, PubMed IDs), provides the alert the readers need, and which they will nevertheless always have to verify.

¹⁰ A venue dedicated to post-publication peer-review.

¹¹ A database dedicated to displaying the polarity of citations (*supporting*, *mentioning*, *contrasting*), with a specific method to detect editorial notices (Uppala et al., 2022).

6. Conclusion

This study is both an update of previous studies about the chemical literature and the preliminary work required to prepare the investigation on other journals and in other disciplines. At this stage, we can but agree with the conclusions of Garfield (1991) and Hubbard (2010) who, at different period of times, urged for a better indexing of corrections, a better linking of correction notices to the original publication and also for empowering users and educating them to identify and locate potential corrections (e.g., to ensure that errors are not included in systemic reviews (Royle & Waugh, 2004)). This should be part of research integrity training programs.

The integrity of the scientific record can only be improved by the joint action of all the stakeholders in the community (publishers, researchers, librarians and the whole range of bibliographic information providers). This kind of study is therefore essential to mobilize them all. In the context of this STI 2023 conference, it is also an opportunity to remind those in charge of scientometric studies or research evaluation procedures to be aware of the shortcomings of the bibliographic databases they rely on it for their analyses.

This preliminary work will need to be complemented by further studies on a broader scope. These studies should also take into account other manifestations of disagreement in science, to be considered as corrective actions, which may include Comments or Letters to the Editor, post-publication peer-review comments or critical citations.

Open science practices

The dataset is available online:

Bordignon, Frederique, 2023, "JACS correction notices - Dataset (2000-2023)", DOI: 10.57745/W7FB4W

Competing interests

No competing interest

Funding information

The NanoBubbles project has received Synergy grant funding from the European Research Council (ERC), within the European Union's Horizon 2020 program, grant agreement no. 951393.

7. Bibliographic references

- Addelston, Aaron., & Goldsmith, U. J. (1966). Procedures for Detecting Errors in Chemical Literature. *Journal of Chemical Documentation*, 6(3), 126–129. https://doi.org/10.1021/c160022a002
- Bordignon, F. (2023). *JACS correction notices—Dataset (2000-2023)* [Data set]. Recherche Data Gouv. https://doi.org/10.57745/W7FB4W
- Bouter, L. (2023). Research misconduct and questionable research practices form a continuum. *Accountability in Research*, 1–5. https://doi.org/10.1080/08989621.2023.2185141
- Casadevall, A., Steen, R. G., & Fang, F. C. (2014). Sources of error in the retracted scientific literature. *The FASEB Journal*, 28(9), 3847–3855. https://doi.org/10.1096/fj.14-256735

- Cooper, E. R. (1992). Identifying errata and retractions: Simplified approaches for serials management. *Serials Review*, 18(4), 17–20. https://doi.org/10.1016/0098-7913(92)90066-6
- Coudert, F.-X. (2019). Correcting the Scientific Record: Retraction Practices in Chemistry and Materials Science. *Chemistry of Materials*, 31(10), 3593–3598. https://doi.org/10.1021/acs.chemmater.9b00897
- Dougherty, M. V. (2019). Correcting the scholarly record for research integrity. *Monash Bioethics Review*, 37(1–2), 76–78. https://doi.org/10.1007/s40592-019-00093-x
- Garfield, E. (1991). How to Avoid Spreading Error Scientists Must Search for Corrections. *The Scientist Magazine*. https://www.the-scientist.com/commentary/how-to-avoid-spreading-error-scientists-must-search-for-corrections-63583
- Hendricks, G., Lammey, R., & Rittman, M. (2022). *Towards a connected and dynamic scholarly record of updates, corrections, and retractions*. MetaArXiv. https://doi.org/10.31222/osf.io/6z7s3
- Hinchliffe, L. J. (2022). The version of record as a central organizing concept in scholarly publishing. *Information Services & Use*, 42(3–4), 309–314. https://doi.org/10.3233/ISU-220164
- Hu, G., & Xu, S. (Brian). (2020). Agency and responsibility: A linguistic analysis of culpable acts in retraction notices. *Lingua*, 102954. https://doi.org/10.1016/j.lingua.2020.102954
- Hubbard, D. E. (2010). Corrections in the Chemical Literature: Their Number and Nature. *Science & Technology Libraries*, 29(1–2), 130–143. https://doi.org/10.1080/01942620903579427
- Jones, D. A., Watson, M. M., Comegys, M., Burnett, A., & Tucker, B. (2003). Errata and Retractions in Electronic Journals. *Journal of Hospital Librarianship*, 3(2), 19–27. https://doi.org/10.1300/J186v03n02_02
- Kiang, N. Y. (1995). How are scientific corrections made? *Science and Engineering Ethics*, *1*(4), 347–356. https://doi.org/10.1007/BF02583252
- Lindberg, D. A. B. (1987). Retraction of Research Findings. *Science*, *235*(4794), 1308–1308. https://doi.org/10.1126/science.235.4794.1308.b
- NISO. (2008). NISO RP-8-2008, Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group. NISO. https://doi.org/10.3789/niso-rp-8-2008
- Noel, M. (2020). Back to disciplines: Exploring the stability of publication regimes in chemistry: the case of the Journal of the American Chemical Society (1879–2010). *Humanities and Social Sciences Communications*, 7(1), 57. https://doi.org/10.1057/s41599-020-00543-6
- Poworoznek, E. L. (2003). Linking of errata: Current practices in online physical sciences journals. *Journal of the American Society for Information Science and Technology*, 54(12), 1153–1159. https://doi.org/10.1002/asi.10320
- Royle, P., & Waugh, N. (2004). Should systematic reviews include searches for published errata? *Health Information & Libraries Journal*, 21(1), 14–20. https://doi.org/10.1111/j.1471-1842.2004.00459.x
- Sabine, J. R. (1985). The Error Rate in Biological Publication: A Preliminary Survey. *Science, Technology, & Human Values, 10*(1), 62–69. https://doi.org/10.1177/016224398501000108
- Teixeira da Silva, J. A., & Dobránszki, J. (2017). Notices and Policies for Retractions, Expressions of Concern, Errata and Corrigenda: Their Importance, Content, and Context. *Science and Engineering Ethics*, 23(2), 521–554. https://doi.org/10.1007/s11948-016-9769-y

- Teixeira da Silva, J. A., & Vuong, Q.-H. (2022). Fortification of retraction notices to improve their transparency and usefulness. *Learned Publishing*, *35*(2), 292–299. https://doi.org/10.1002/leap.1409
- The ACS Guide to Scholarly Communication. (2020, January). https://doi.org/10.1021/acsguide 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. https://doi.org/10.48550/arxiv.2210.09553