

Improving the quality of open-access submissions

UKSG eNews 508

4 February 2022



Michael Upshall

UNSILO, Cactus Communications

Scholarly publishing, whether published open-access or by subscription, has an enviable record of quality. Given the number of new articles published every year – estimates range between three and four million – the number of articles retracted or corrected is tiny, no more than 500 or 600 (according to [Retraction Watch](#)). Nonetheless, given the steady increase in the number of articles published, it is no longer possible to carry out all the standard checks on a manuscript that were traditionally done in the publisher's offices by hand. The age of manual checking is over; but on that point, as far as the manuscript submission process is concerned, open-access publishing has a lesson for traditional publishers can learn from.

Several companies, including UNSILO, **Ripeta**, **Scholarcy**, and others, have in recent years focused on identifying checks to aid the submission process, by examining the submission workflow and discussing it with authors and publishers. In this way, well over 50 separate operations have been reviewed that can identify potential problems with an article – and many of these can use automation to help. With the seemingly inexorable rise of articles published, it seems obvious that the use of technology to assist in running a standard set of checks to ensure articles are of satisfactory quality is essential. Given the time constraints publishers work with, making the human publisher's time more effective seems an essential task.

As far as the submission process is concerned, Open Access and subscription journal publishing should be identical. The author writes an article, and checks it for a number of criteria, then the publisher also checks it before acceptance for publication. It is surprising, therefore, to see how much of a difference there is between submission workflows in fully OA publishers and those of the more traditional publishers. While in both cases the author and the publisher share a common goal in getting to publication as quickly as possible with

the highest possible quality, OA-only publishers have typically put more focus on using automation to assist the workflow than traditional publishers. Perhaps this is the inevitable consequence of setting up business processes: once a business process is set up, there is no incentive to change what already works. The large traditional publishers have been processing manuscript submissions for many years, while the new OA publishers, in contrast, setting up a system from scratch, did not need to adhere to traditional practices. As a result, OA publishers tend to be using more automated checks than subscription publishers.

As an example of the kind of tools available, consider the checks available to authors and to publishers. Several companies provide suites of checks for authors to run on their own manuscript before submission, checking bibliographic style, count of words, and so on. In addition, there are several checks a publisher would want to check independently of the author: broadly speaking, these cover the integrity of the publication. These include checks for plagiarism, for image manipulation, for salami-slicing, for excessive use of self-citations, among others. Recent initiatives are also assessing any data provided as part of the article.

What makes an effective automatic check? Broadly speaking there are two types. The simplest check is just finding or counting. Counting the number of words in a submission is trivial, but far quicker than doing it by hand. Identifying references at the end of the article that have not been cited in the article is still a very simple check, but something that would take a human many minutes to carry out. Even if 95% of submissions are perfect in this respect, to run an automatic check ensures a more consistent quality.

More advanced are checks using a training set. For example, UNSILO has a check for conflict of interest statements in a manuscript. This check does not simply search for phrases including “conflict of interest” or similar syntactic variants, but makes use of a training set compiled by humans reading hundreds of similar articles. Whenever a human reader sees a sentence in the training set that appears to have a possible conflict-of-interest implication, the sentence is highlighted for the author (or editor) to review; an example might be “this article was sponsored by XYZ Labs”, or “ABC Pharmaceuticals commissioned the research”. Importantly, in a check of this kind, the machine does not make any judgement. All it does is find potential problems, and then provides the evidence for the human to make a decision. The process resembles the combination of machine- and human checks on fraudulent credit card transactions. There are far too many transactions for any human team to assess by hand, so a machine identifies unusual transactions for a specific kind of account, and a human examines the queried transactions to check if they should be investigated further or not. The result is a more secure process, for the benefit of the user and the bank.

AI tools can be used with great effectiveness to check the integrity of a publication. Some of the checks available include:

- Excessive self-citing
- Substantial overlap with other articles on the same topic: this is traditionally done using string matching, but more advanced systems use concept matching to identify if the same argument is being presented using different phrasing
- An unusually high number of old references

The effective use of AI-based tools – to identify, not necessarily to fix problems – can ensure a consistently higher quality of submission without reducing time to publication. As always with AI, the effective solution is not to rely on the AI to deliver solutions, but to provide evidence to the human so they can make an informed decision. While most of the major open-access publishers make use of these tools already, traditional subscription publishers can make use of the same approach.

These views are the authors' own and do not necessarily reflect the views of UKSG.



This UKSG Editorial is taken from the industry newsletter UKSG eNews, published every two weeks exclusively for UKSG members. The newsletter provides up-to-the-minute news of current issues and developments within the global knowledge community.

To enjoy *UKSG eNews* and other member benefits [become a UKSG member.](#)