Delays in Open Access Availability of Faculty Journal Articles

John Wiswell, Appalachian State University

How much Open Access is there?

Some global estimates of Open Access (OA) availability of published research have been surprisingly high. Archambault et al, estimated that of 2013 articles, nearly 50% were OA in 2014 (2014, p.35). Adjusting for the undercounting, they actually estimate about 55% of articles from the year before were available OA. This estimate includes Europe (and the entire world), which may be higher than in the United States.

32% in the USA?
There are a few estimates of OA availability at more local levels. Swan et al, give an estimate for one US university, MIT, that 32% of 2011-2013 articles were available OA (2015, p.27). They also estimate that the average delay in OA availability for Duke University-generated articles is now just 5 months (p.30). Duke is one of the universities with the lowest average delay.

My own experience
My own experience trying to access articles is that the proportion is below 20%, much lower than the estimates. Also, one of my responsibilities is to recruit faculty articles for our institutional repository, while dealing with the difficulty of sustaining the workflow and meeting publisher restrictions.

A few hypotheses, none of which will be tested
This study will not involve hypothesis testing, but I have had several hypotheses, not all of which will be addressed in the brief interval in this phase of this study:

1. The OA proportion of recently published articles would be less than 20%.
2. The proportion varies across universities, departments, and disciplines.
3. The proportion will vary over time.
4. Talking about it will encourage participation.

Methods

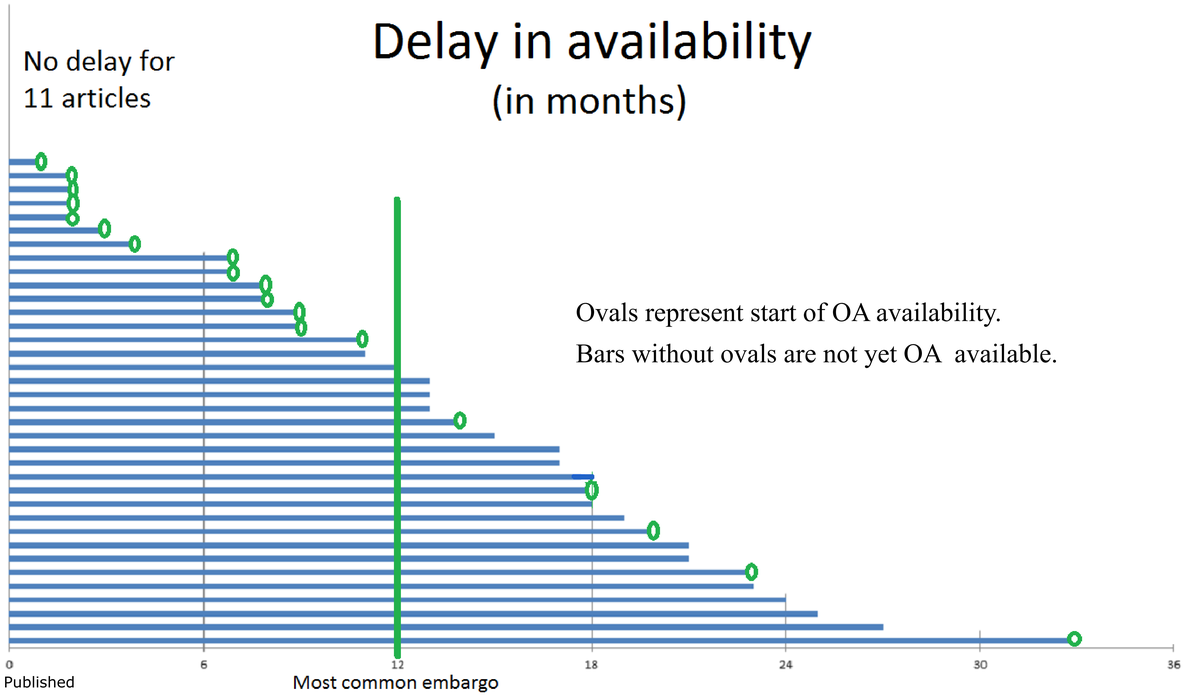
Finding articles from 12 departments
I searched several databases and search engines (Web of Science, PubMed, Google Scholar, EBSCO Discovery, and ProQuest databases) for all refereed journal articles, published in 2014 from 12 selected departments at masters-comprehensive, regional, public universities. The selected departments include a range of natural science, social science, and health science disciplines. None were from the arts and humanities. I identified 144 articles from those departments.

Creating a sample and clean up
A random sample of 50 articles was taken from the 144 identified articles. Two of these were removed when it was found that they were not peer-reviewed, and a third article was removed, because it had not yet been assigned to an issue dated 2014. The sample then included 47 articles.

Variables
Data were found on the 47 articles remaining in the sample, including month that they were nominally published, month that they became available by subscription, and month they became available in a reasonably findable Open Access website. Journals were searched in the Sherpa Romeo database for information on Open Access restrictions, especially embargo periods for posting post-peer-review versions of the articles. The intervals from nominal and actual publication to open access availability were calculated.

Analysis
The sample is small enough that they can be visualized easily. Descriptive statistics could not be meaningfully calculated, since a large minority of the sample is not yet available Open Access. In statistics-speak, those observations are censored. The data will simply be displayed instead.

Probability of continued delay is similar to survival and failure data, so it can be also be displayed as a Kaplan-Meier plot. However, this is mostly for fun. There are enough articles for which delay continues that the probabilities are (probably) underestimated.



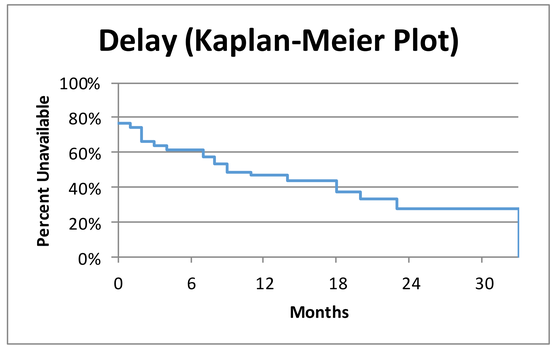
Months to full text

Blue is OA. Red is not yet OA.

0 0 0 0 0 0 0 0 0 0 0 0
1 2 2 2 2 3 4
7 7 8 8 9 9 11
11 12 13 13 13 14 15
17 17 18 18 18 19 20 21 21 23
23 24 25 27 33

Probability of continued delay

This is a statistical method for "survival" data, but not very helpful.



Results

The analysis was done in early October 2015. The articles all have nominal publication dates in 2014, but actual publication online started in 2012 and 2013 for some.

- 11 of the 47 study articles were available Open Access in the same month they were published or made available online. That's 23%.
- 14 (30%) of the remainder were made available OA in the first 11 months. That's a total of 53% before a year passes.
- 5 (11%) were made available between 12 and 33 months. That's 64% now available.
- 17 (36%) are still not available Open Access.

Most of the articles were published by fairly large publishers, and a few (4) were published in Open Access journals, that is, Gold OA journals.

When the full text articles were available, there was a range of websites where they had been deposited: personal, universities and departments, government, scholarly associations, and others. Often collaborating authors at non-study institutions appear to have made them available. It is not apparent that other actors, apart from researchers and their institutions, have played an important role in availability.

Even with the Sherpa Romeo database, it is often not clear what the applicable publisher restrictions and permissions are. The 12-month embargo for post-prints is very common.

Conclusions

The researcher's expectation was that there would be substantial underuse of Open Access opportunities, and that has been the case. This is an area for work in communicating with faculty and in negotiating barriers. However, the naïve researcher has been surprised by the heavy availability of full-text articles at intervals shorter than publishers' apparent embargo periods. This might be another area for further research and discussion between librarians and researchers.

In its way, the 12-month embargo appears to be fairly effective in decreasing Open Access availability. The evidence here shows that few articles (5 of 22) are made available after 12 months and the delays are substantial and continuing. On the other hand, publisher restrictions appear possibly to speed up availability of other articles.

This study has several limitations. The departments sampled are not necessarily representative of the region's publishing, even from the natural, social, and health sciences. The observations are not independent, and the results may be biased by individual, departmental, university, and disciplinary practices. (However, it does not appear that any of the sampled departments were under any Open Access policies during this period.) The sample was randomly chosen, but the population consists of publications from only 12 departments.

These results suggest that further study would be valuable. A more systematic sample of more units at more institutions might give more generalizable results. Also, qualitative approaches to studying researcher behavior and preferences could be valuable.

References

Archambault, Didier Amyot. 2014. "Evolution of Open Access Policies and Availability, 1996-2013." Science-Metrix.

Swan, Alma, Yassine Gargouri, Megan Hunt, and Stevan Harnad. 2015. "Open Access Policy: Numbers, Analysis, Effectiveness." arXiv Preprint arXiv:1504.02261.

Contact

John Wiswell, Belk Library, Appalachian State University, wiswellj@appstate.edu, (828)262-7853