Gold, Green and Black Open Access

Bo-Christer Björk

Hanken School of Economics, Helsinki, Finland

This is a preprint version (prior to peer review) of an article published in Learned Publishing, electronically published 7th February 2017, DOI: 10.1002/leap.1096. The changes after peer review were minor. If you site this article, please site the published version

Key Points

- The debate about open access has until now focused on the gold (journals) versus the green route (manuscript self-archival), as the main alternatives
- Recently an even more disruptive form of OA has emerged, in the form of illegal article copies retrievable from academic social networks or pirate sites.
- This “black open access” provides access to a large part of the pay-walled article output which cannot be found in repositories.
- While Black OA don’t yet threaten the revenues of major subscription publishers, its rapid emergence is a further symptom of the dysfunctionality of the current subscription system.

Introduction

Universal Open access to scholarly research publications is deceptively simple as a concept. Any scientific publications, whether found via a Google keyword search, or by trying to access a citation would be just one click away. But the path to get there from the current subscription dominated journal publishing model has proved to be complex and filled with obstacles. Since the terms gold and green OA were coined almost fifteen years ago, much of the debate inside the Open access movement has been focused on the relative merits of these two paths [1].

Currently gold OA seems to be growing in importance: There are a number of factors contributing to this. Firstly, the emergence of successful professional OA publishers and megajournals offering rapid publication via an innovative peer review approach. The second coming of hybrid OA publication is also of increasing impact, fuelled by the new APC funding mechanisms put in place my major research funders like the Wellcome Trust and the Research Councils UK. New kinds of nationwide super-bundle licenses that cover both subscription access and hybrid charges, could further establish hybrid as a conversion vehicle to full OA [2].

Traditional green OA in institutional repositories has been struggling with getting researchers to upload, despite the fact that most major universities now have such repositories in place [3]. The researchers just don't seem bother with the little extra work involved and many are ignorant of the possibilities. The leading subject repositories, arXiv and PMC are doing better, but only cover some fields of science. And publishers have tightened embargo rules for self-archiving, making green OA less attractive.

All in all, the progress towards full OA has been slower than hoped for. Currently somewhere between 35-50 % of all research articles from recent years can be found as either gold, hybrid or self-archived repository copies. Empirical studies have shown varying shares depending on the definitions and methods employed [4, 5, 6]. Some studies also identify articles as OA, which publishers make open for promotional purposes. It is quite common for a publisher to have the first issue of the current year open, but this also means that the articles may disappear behind a pay-wall when the year changes.

A recent development muddling the picture is the emergence of other channels for uploading and accessing research articles without subscriptions, payments and bureaucracy. In contrast to proper repositories which by and large enforce the licence rules, these new channels mostly offer the exact digital replicas of the published journal articles, and do so illegally. The two main players are academic social media, especially Research Gate, and pirate copy web sites such as Sci-Hub. I propose to call this black OA. Black as in pirate! I was first considering the milder colour grey, but the term grey literature has already an established meaning in the context of scholarly publishing covering theses, government reports etc. [7]. Grey literature has in fact greatly benefitted from OA. The term Platinum OA [8] has also been proposed for non APC-charging gold OA (free also for authors), but the term has never really caught on.

Academic Social Networks

Academic Social Networks (i.e. Research Gate, Mendelay, Academia.edu) have their origin in the tremendous success of networks like Facebook and LinkedIn. Like for all such networks the key word is critical mass. If enough colleagues of yours have already joined the incentive increases.

Some ASN's also have a dubious practice of spamming non-members with emails seemingly coming from colleagues, which actually are sent automatically by the systems. ASNs also try to offer other features like citation counts of their own, user profiles et.c.[9].

The leading ASNs are all business ventures where the revenue model, since they don't charge members, is to build up a big enough user base to exit via being bought by a bigger company [10]. The prime case is Mendeley, which was purchased by Elsevier in 2013. Elsevier also bought the much older SSRN (Social Sciences Research Network) in 2016.

Although most the articles uploaded to ASNs are direct PDF copies of the published articles, its perfectly possible to upload personal versions of the manuscripts. But even this might not be in accordance with many license agreements that authors have signed with publishers.

So far massive law suites against ASNs have not started, but for instance Elsevier has sent out takedown requests to Academia.edu [11].

Pirate copy sites

The role model for sites of pirate copies of scholarly journal articles is Napster. Here the driving force is not financial gain but indignation that a few big publishers are making big profits "off the back" of the global academic community. The notion that the results of research is a public good which should be free is common with advocates of gold and green open access, but the approach is more radical in actually breaking the law. To date Sci-Hub is the notable example, although there have been earlier incidents of massive downloading of scholarly articles by OA activists [12]. Another related way of obtaining article copies is to twitter messages including an article title (i.e. #icanhazPDF) asking unknown academics who have subscription access to download a copy and mail it to the tweeter [13]. Obviously a much stronger case can be made for pirating scholarly articles than music or films, since the authors don't lose any income by this action.

As blatantly illegal as Sci-Hub for instance is, its been difficult for the publishes to take legal action, due to the fact that the founder and manager is based in Russia. What they have achieved so far is that Sci-Hub has been forced to cede some domain names, but the service has so far managed to quickly resurface using slightly altered ones [14]. As for approaches like #icanhazPDF, although the actions are in principle illegal, it would be extremely difficult to take legal action.

Why has black OA become so popular

Black OA has grown very rapidly in the last three to four years. The reasons are slightly different for the two subspecies. Authors who have joined Academic Social Networks probably find it easy to just post the exact published pdf of their article there, no questions asked. The alternative, using their institutional repository, would entail first finding out the exact meaning of the complex copyright and license agreements they sign with the publisher. After that there would often be the extra work of formatting the manuscript, inserting a title page, and reformatting for instance tables and figures which often have been separate files before submission.

From the reader's viewpoint Academic Social Network copies are as easily findable as other types of green OA. In a google scholar search a free Research Gate copy will show up in a column to the right, the same way an institutional or subject repository one will. A basic google search will

primarily point to the publisher's version, usually behind a pay wall. But simply adding "Research gate" to the title in the search string will take you to the copy, if there is one.

As for Sci-Hub it's a pure pirate site, which hasn't asked for consent from neither authors nor publishers. The repository has downloaded more than 50 million journal articles directly from the publishers' web sites [15]. This has been accomplished using the passwords of individual researchers to their university library portals for mass downloading, but it is unclear if the passwords have been voluntarily given by the individuals in question, or whether they have been obtained by phishing.

Slightly cumbersome for readers is that Sci-Hub copies don't show up in Google or Google Scholar searches. The user has to go the site itself (which is constantly changing name, due to lawsuits from the big publishers) and input the DOI or the title of the article he searches for. In case it can't be found this actually triggers Sci-Hub software to go and find and retrieve it. Hence the site is used more as a last resort when a free copy cannot be found otherwise. It's definitely quicker than mailing authors in the hope of obtaining a copy, not to mention the archaic and slow method of interlibrary loan, if that is even a possibility. And for many younger academics and students paying pay-per-view charges is out of the questions. The strength of Sci-Hub is currently that its much more comprehensive than ASN:s or repositories which only contain selected articles rather randomly. And many of the users of Sci-Hub, despite this say they use it more for convenience than to save cost [16].

In summary for both ASNs and pirate sites there are three major reasons why they have become so popular with both authors and readers. Easy of use, the fact that they are found morally acceptable by most academics [17] and the almost non-existent legal risks involved in up- and downloading individual articles.

Conclusions

In a nutshell around 25 % of recent research articles (after a year) are accessible as gold OA, and an additional 15-20 % can be found as legal green copies. Of the remaining 50 % the vast majority are findable as illegal black OA copies.

The direct negative effects of black OA on mainstream subscription publishing, in terms of decreasing revenue, have so far probably been negligible. A major reason is that most of the sales revenue is in the form of huge electronic big deals, which usually span several years. Also the publishers exert huge lock-in power, so that it is almost impossible to a library refuse a deal with one of the major publishers. There is thus a lot of inertia in the system isolating it from negative impacts of both green and black OA. The loss of income could potentially be bigger in less developed countries with strained university budgets, but publishers get much less income from those anyway and in many cases offer free access.

The biggest effect of black OA could in fact be in diluting the popularity of green OA channels, in combination with publishers tightening the embargo rules which in particular institutional repositories tend to follow. This is parallel to the adverse effects that predatory OA publishing has on authors willingness to submit to serious gold OA journals, due to the tainting of image of author paid OA.

If anything black Open Access should be seen of a symptom of a seriously flawed mainstream scholarly publishing model, which has failed to adequately adapt to the needs of the international

research community. Black OA is obviously not a solution to the problems, the remedy is in my opinion 100 % gold OA for research papers funded by APCs or other mechanism. When we get there the “market demand” for retrieval via sites like Research Gate and Sci-Hub disappears, and the use of subject and institutional repositories will be reduced to preprint manuscripts, grey literature and overlay purposes.

Acknowledgements

On a personal note I have a wide-ranging experience of the channels discussed in this article. I’ve published in both free and APC-charging OA journals, in a delayed OA journal, and some articles of mine in subscription journals are hybrid OA. I’ve uploaded many manuscript copies to my university’s institutional repository, to a personal website shared with a colleague and some of my articles can be found in PMC. Co-authors have uploaded articles to Research Gate, and the one subscription journal article I specifically checked via Sci-Hub was indeed retrievable.

References

1. Harnad, S., Brody, T., Vallières, F., Carr, L., Hitchcock, S., Gingras, Y., Oppenheim, C., S, Stamerjohanns, H., & Hilf, E. R (2004) The Access/Impact Problem and the Green and Gold Roads to Open Access, Serials Review, 30:4, 310-314.
2. Björk, B-C. (2016) The open access movement at a crossroads – are the big publishers and academic social media taking over? 29: 2, 131–134, doi: 10.1002/leap.1021
3. Eisen, M. (2015) The inevitable failure of parasitic green open access. Blog post, 25.5.2015. http://www.michaeleisen.org/blog/?p=1710
4. Archambault, E., Amyot, D., Deschamps, P.,Nicol. A, Rebout, L., & Roberge, G. (2013) Proportion of Open Access Peer-Reviewed Papers at the European and World Levels—2004-2011, report http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf
5. Björk, B-C., Laakso, M., Welling, P., Paetau, P., (2014). Anatomy of Green Open Access, Journal of the Association for Information Science and Technology, 65:2, 237-250.
6. Research Information Network (2015) Monitoring the Transition to Open Access, A report for the Universities UK Open Access Co-ordination Group, https://www.acu.ac.uk/research-information-network/monitoring-transition-to-open-access
7. Schöpfel, J. 2011 Towards a Prague Definition of Grey Literature, Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues. Prague, 6-7 December 2010, pp. 11-26, https://archivesic.ccsd.cnrs.fr/sic_00581570/file/GL_12_Schopfel_v5.2.pdf
8. Haschak, P. (2007). The 'platinum route' to open access: a case study of E-JASL: The Electronic Journal of Academic and Special Librarianship, Information Research, 12(4) paper 321. http://InformationR.net/ir/12-4/paper321.html

9. Nicholas, D., Clark, D. and Herman, E. (2016) ResearchGate: Reputation uncovered, Learned Publishing, 29:3, 173–182, DOI: 10.1002/leap.1035
10. Matthews, D. (2016). Do academic social networks share academics' interests?. Times Higher Education. https://www.timeshighereducation.com/features/do-academic-social-networks-share-academics-interests
11. Howard, J. (2013) Posting Your Latest Article? You Might Have to Take It Down, The Chronicle of Higher Education, 6.12.2013, http://www.chronicle.com/blogs/wiredcampus/posting-your-latest-article-you-might-have-to-take-it-down/48865
12. Murphy, S. (2011) ”Guerilla Activist” Releases 18,000 Scientific Papers, MIT Technology Review, 22.7.2011, https://www.technologyreview.com/s/424780/guerilla-activist-releases-18000-scientific-papers/
13. Gardner, G. and Gardner, C. (2015) Bypassing Interlibrary Loan Via Twitter: An Exploration of #icanhazpdfRequests, American Library Association, ACRL 2015 Proceedings, http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/2015/Gardner.pdf
14. Kravets, D. (2016). "Piracy site for academic journals playing game of domain-name Whac-A-Mole". Ars Technica. http://arstechnica.com/tech-policy/2016/05/piracy-site-for-academic-journals-playing-game-of-domain-name-whac-a-mole/
15. Bohannon, J. (2016) Who’s downloading pirated papers? Everyone. Science, 26.4.2016,DOI: 10.1126/science.aaf5664
16. Charleston Conference (2016) Who's Faster, a Pirate or a Librarian? Sci-Hub, #icanhazPDF, Alternative Access Methods,and the Lessons for Librariesand their Patrons, Panel session, 2016 Charleston Conference; Issues in book and Serial Acquisition, 31.10 - 5.11. 2016, Charleston, USA, https://2016charlestonconference.sched.org/event/85ez/whos-faster-a-pirate-or-a-librarian-sci-hub-icanhazpdf-alternative-access-methods-and-the-lessons-for-libraries-and-their-patrons
17. Travis, J. (2016) In survey, most give thumbs-up to pirated papers, Science, 6.5.2016, DOI: 10.1126/science.aaf5704