

Removing all barriers in the way of Science: A case study on Sci – Hub’s usage in Europe

Gerakopoulou Elli¹, Founti Maria Christina² and Foster Christopher³

¹ The University of Sheffield, Sheffield, UK elli.gerakopoulou@gmail.com

² The University of Sheffield, Sheffield, UK cmfounti1@sheffield.ac.uk

³ The University of Sheffield, Sheffield, UK christopher.foster@sheffield.ac.uk

Abstract. Sci- Hub is the world’s largest pirate academic library. This case study explores its usage in Europe based on the website’s log data for the 6-month period between September 2015 and February 2016. A quantitative analysis of the log data was conducted followed by a qualitative interpretation of the results. The analysis revealed the European countries with the largest number of download requests, emerging download trends across disciplines and academic publishers. Financial crisis impact, academic publishing regimes and user behavior studies are used to explain the results. In addition, hacktivism and Open Access are explored and suggest Sci-Hub’s alignment with hacktivism and the formation of a unique digital social movement.

Keywords: Sci – Hub, Piracy, Open Access, Hacktivism, Europe.

1 Background

Sci – Hub is the world’s largest digital pirate library that makes scholarly papers freely available, by-passing publishers’ paywalls. It was created in 2011 by Alexandra Elbakyan, a Kazakh neuroscientist and rapidly became popular, building a large collection through automated processes and simple access. Sci – Hub’s ethos for providing access to knowledge is embedded within wider discourses of hacktivism and an activist critique of academic publishing regimes. It thus provides an important case to explore academic publishing initiatives in Europe.

2 Literature

In recent years, we have seen the growing commercialization of academic publications with mergers amongst large publishers. In 2013, five out of around 2,000 academic publishers were responsible for half of the total publications [9]. Moreover, from an economic perspective, scientific research papers can be seen as a ‘natural

monopoly' limiting typical competitive drivers in the sector [13]. With academics being increasingly result-orientated and pressured to publish within well-established high impact journals [10], academic publishing is moving towards an oligopoly.

As a consequence, academic libraries are facing increasing financial pressures. Library subscriptions are becoming unaffordable, with academic library subscription fees having increased by 145% the last six years [12]. With static or reducing funds, even the best academic institutions are facing a 'serials crisis', struggling to keep up with the latest volumes [5].

Responding to these challenges, open access (OA) is the foundation for providing fair access to scientific knowledge, and OA principals have been subscribed to by government and academics [2]. Practical OA strategies in practice include so-called 'Gold OA' where material is fully open to readers in journals, and 'Green OA' where academics self-archive material. Elsewhere, funders are beginning to mandate that public research is disseminated in an OA-compliant form [13]. Such OA strategies have expanded the availability of academic publishing. Recent research suggests that Gold OA content is now almost 20 per cent of total journal content [7]. OA journals also appear to be expanding with a 30 percent increase in the number of articles since 2000 [8].

OA is growing. However, growth is at a relatively modest rate. Moreover, critics have argued that OA strategies marginalize the voice of scientific producers and consumers. OA negotiations resolve around political negotiations, where academic with greatest stake in disseminating knowledge are marginalized. OA practice often folds into atomized activity, where OA is dependent on individualized archiving.

With their lax attitude to copyright, pirate libraries have often been dismissed as the work of individual 'hackers' and well outside the mainstream. We argue this position is problematic, it underplays the coherent political agendas, and new forms of social movement through 'hactivism' that are central to political agency in the digital society [3]. Pirate libraries align with the movement of 'Guerilla OA' that emerged following the ordeals of the hacker-activist Aaron Schwatz. Through social networks, popular academic hashtags such as #ICanHazPdf and #PdfTribute emerged, integrating public protest by individuals with the gift economy. Pirate libraries such as LibGen and Sci-hub have formed with overt political goals around access to knowledge, their approaches akin to civil disobedience that have been effective part of social movement for centuries [4].

A limited number of empirical studies have been done on pirate libraries and they have mainly lacked a clear analytical position [6]. Drawing on hacktivist ideas this work explores pirate libraries, raising questions about digital divide, intellectual property and ethical issues. The aim of this preliminary study is to use the recent log data released by Sci-hub to provide a better understanding of use across Europe, exploring links to economic factors and user behavior.

3 Methodology

The analysis was conducted using the Sci-hub dataset shared by Bohannon [1]. The available log data corresponded to a total of 28 million download requests over the six months from September 2015 to February 2016. The log data consist of the DOI of each paper, the timestamp and anonymized location of users.

Data analysis was facilitated through KNIME big data analytics platform, which is useful for creating workflows and easily filtered results. All designed workflows were validated against Bohannon's [1] results to confirm for triangulation. Additional data were also obtained from the Crossref which provides details of the 9,056 DOI prefixes. The results were categorized using Plum Analytics subjects [11].

4 Preliminary Findings

About a quarter (22, 3%) of the total Sci- Hub download requests were made from European countries. This study categorized ten countries with the highest number of download requests; the Mediterranean (Portugal, Spain, Italy, Greece), Eastern European (Ukraine, Poland), and central European (France, Germany, UK, the Netherlands).

The publishers with the highest requests were mainstream publishers, such as Elsevier and Springer-Verlag. About 70 per cent of the downloads corresponded to publishers among the top ten. Lastly, the majority of the 20 most downloaded items came from the hard sciences and corresponded to more than half of the downloads. Nearly all of these papers were in the overall top downloaded papers in Plum Analytics data. From these, 80 per cent were available as open access or free from publishers.

5 Discussion

Based on the results, the country groups highlight a connection between regions reducing funding to academic libraries and the use of Sci-hub. Financial crisis that started in 2008 has resulted in funding restrictions, impacting higher education. This has particularly been felt in Eastern European and Mediterranean countries which are disproportionately represented in the results. Sci-hub represents an alternative for users as academic libraries struggle to stay up-to-date with subscriptions.

Findings illustrate that most download requests, corresponded to publications by a small number of academic publishers. These results show a higher skew towards large publishers than would be expected from recent estimates in the literature. A likely explanation is that the subscription demands of large publishers are a strong driving force for users to Sci-hub.

The findings show that publications from the hard sciences had the highest number of downloads which align with previous studies. Although the majority of the scholarly items among the top 20 are available in OA, this had little impact on downloads.

These findings are an indication that Sci-hub is not being used by niche fields but is rather used broadly across fields. The open access findings potentially highlight the weaknesses of paper-by-paper OA and the complications of users finding and accessing relevant OA material compared to the simplicity of Sci-hub.

6 Conclusions

These early findings suggest Sci-Hub's impacts broadly align with the goals of hacktivists and OA activists. Sci-hub use is proportionally higher in regions that are being hit hardest by access costs. Downloads are also more skewed towards those oligopoly publishers that have been subject to critique. Findings on OA availability appear to align with activist critiques on limitations of individualist approaches to OA. Further cross-referencing is likely to yield further explanation as we examine this data.

References

1. Bohannon, J.: Who's downloading pirated papers? Everyone. *Science* 352(6285): 508- 512 (2016).
2. Budapest Open Access Initiative, <http://www.budapestopenaccessinitiative.org>, last accessed 2017/05/17
3. Coleman, G.: Hacker, hoaxer, whistleblower, spy: the many faces of Anonymous. Verso, London (2014).
4. Crossley, N.: Making sense of social movements. Open University Press, Philadelphia (2002).
5. Dingley, B.: U.S. Periodical Prices – 2005. U.S. Periodical Price Index 2005, pp 1-16. Association for Library Collections and Technical Services (2007).
6. Gardner, G.J., McLaughlin, S.R., & Asher, A.D.: Shadow libraries and you: Sci-Hub usage and the future of ILL. In: ACRL 2017. Association of Research and College Libraries, Baltimore (2017). <http://eprints.rclis.org/30981/>, last accessed 2017/07/19
7. Kramer, D.: Steady, strong growth is expected for open-access journals. *Physics Today*, 70 (5): 24 (2017).
8. Laakso, M., Welling, P., Bukvova, H., Nyman, L., Björk, B.C., & Hedlund, T.: The development of open access journal publishing from 1993 to 2009. *PLoS ONE*, 6(6): e20961 (2011).
9. Larivière, V., Haustein, S., & Mongeon, P. : The oligopoly of academic publishers in the digital era. *PloS ONE*, 10(6): e0127502 (2015).
10. McGuigan, G.S., & Russell, R.D.:The business of academic publishing: A strategic analysis of the academic journal publishing industry and its impact on the future of scholarly publishing. *Electron J Acad Spec Librarianship* 9: 1-11 (2008). http://southernlibrarianship.icaap.org/content/v09n03/mcguigan_g01.html, last accessed 2017/09/03

11. Parkhill, M.: Sci-Hub: The Academic Cat is Out of the Bag. Plum Analytics (2016). <http://plumanalytics.com/sci-hub-academic-cat-bag-post/>, last accessed 2017/09/02
12. Resnick, B.: Why one woman stole 50 Million Academic Papers – and made them all free to read. In: Vox (2016). <https://www.vox.com/2016/2/17/11024334/sci-hub-free-academic-papers>, last accessed 2017/05/29
13. Suber, P.: Open Access, The Book. MIT PRESS, Massachusetts (2012).