Abstract

This paper seeks to explore the impact of shadow library websites such as Library Genesis and Sci Hub on the scholarly communications ecosystems from the perspective of paper authors. Other stakeholders' stances have already been made quite clear. Publishers are firmly against these databases, which infringe copyright and circumvent paywalls, and they have pursued litigation to shut down shadow libraries. The academic publisher Elsevier has successfully sued Sci Hub multiple times for copyright infringement. Those accessing the articles, especially students, are happy to have access to scholarship that they would not be able to afford otherwise. Some opinions may be deduced by examining from the perspective of authors as researchers, while others can be surfaced by looking through debates and other posts on social media. However, a dedicated research study would beget the most useful information. This paper looks at some of the factors that influence an author's decision to speak on shadow libraries and proposes a few starting points to approach an empirical study on authors.

Introduction

Shadow libraries are unauthorized online collections composed primarily of digital copies of copyright protected material. In the field of scholarly communications, this copyright protected material generally takes the form of published journal articles and monographs. Two of the largest and most well known examples of academic shadow libraries are Library Genesis and SciHub. Shadow libraries surfaced as a response to the financial barriers of accessing scholarly literature, which has always been an issue for scholars outside of major North American and European universities, and has been further exacerbated globally by the serials crisis.

This paper seeks to explore the impact of shadow library websites such as Library Genesis and Sci Hub on the scholarly communications ecosystems from the perspective of paper authors. Other stakeholders' stances have already been made quite clear. Publishers are firmly against these databases, which infringe copyright and circumvent paywalls, and they have pursued litigation to shut down shadow libraries. The academic publisher Elsevier has successfully sued Sci Hub multiple times for copyright infringement. The 2015 lawsuit resulted in an injunction, and Sci Hub's web domain was suspended (Schiermeier, 2015). After another lawsuit in 2017, Elsevier was granted $15,000,000 in damages by a U.S. court (Schiermeier, 2017). Those accessing the articles, especially students, are happy to have access to scholarship that they would not be able to afford otherwise. If someone's institution does not subscribe to an academic journal, they may be asked to pay upwards of $30 USD per article to access research (Bohannon, 2016), and even after paying, there's no guarantee that the articles would actually be relevant to their research. The voice of the authors whose papers are being illegally accessed is conspicuously missing in the discussion about shadow libraries. We can extrapolate what some of their attitudes might be by looking at the other roles that they may play in the scholarly publishing ecosystem, but it is difficult to discern their opinions with certainty without engaging in a focused study. This paper looks at some of the factors that influence an author's decision to speak on shadow libraries and proposes a few starting points to approach a serious study on authors.

Literature Review

Shadow library users

The majority of current research on shadow libraries centers on the users accessing scholarly materials. Some of that research attempts to understand what content is being accessed the most, both in terms of academic disciplines and journal publishers. Research in this area has been facilitated by data released by Sci Hub itself. In 2017, the founder of Sci Hub released a list of all DOIs indexed by the site, and researcher Bastian Greshake cross referenced that information with a previous dataset containing information about downloads (2017). Greshake identified three major trends in what users were accessing. Firstly, users were most interested in recent works, "with ~35% of the downloaded publications being less than 2 years old at the time they are being accessed" (2017, p. 4). Second, articles from chemistry and engineering journals were overrepresented in both the overall collection and the most downloaded articles. Third, despite Sci-Hub having material from around 1700 different publishers, the 9 largest publishers account for around 70% of total works and 80% of downloads (2017, p. 4). This data provides a good starting point to identify authors to interview on the topic of shadow libraries. Scholars in certain STEM fields who are currently active in publishing their work in specific journals are the most likely to have their work impacted by being made available in shadow libraries.

In a feature published in *Science*, John Bohannon, the author who worked with Sci-Hub's founder to release the original dataset about downloads, focused on the geographic locations that Sci-Hub downloads come from (2016). Iran, Egypt, and India were the 3 countries downloading the most content, and many of top download locations closely correlate with places with limited journal access. However, looking at the data also showed that many users download content that they should have institutional access to. In interviews with Sci-Hub users, they reveal that they

prefer Sci-Hub because their university access systems are too troublesome or complex to navigate.

The book *Shadow Libraries: Access to Knowledge in Global Higher Education* is an in depth study of shadow libraries around the world. Through a series of case studies in different countries and organizations, *Shadow Libraries* explores the history of shadow libraries and their role in furthering equitable access to information. The book considers how the culture around information and academics, in conjunction with political and economic factors, influenced the way shadow libraries developed (Karaganis, 2018).

Open Access

Shadow libraries and the Open Access (OA) movement are closely related. They both originated in response to the increasing cost and decreasing accessibility of academic journals. Shadow libraries and OA seek to address the same problem, but they go about it in different ways. Sci-Hub prioritizes getting materials out to the public with no restrictions and no regard for copyright, while Open Access tries to work together with authors and publishers to make materials available to researchers. Despite having similar goals, the differences in approach can put the two movements at odds.

Advocates for OA have expressed criticism towards shadow libraries, alleging that their existence makes it more difficult to obtain legal access. Publishers want to place more restrictions on digital access because they are wary of unauthorized distribution, and they may use losses from piracy as a justification to increase subscription prices. University librarians have also noted situations where publishers revoked an institution's access because of excessive numbers of downloads (Harrison et al., 2018). Some librarians believe that authors can and

should play a more active role in furthering access to scholarly works, indicating that there is interest in what authors think about the issue.

Authors

Information about author perspectives are difficult to find in part because of the illegality of shadow libraries. After a situation in 2019 where Elsevier threatened legal action against the bibliographic management tool Citationsy for linking to Sci-Hub in a blog post titled "Hacking Education: Download Research Papers and Scientific Articles for Free" (McKenzie, 2019). Although Citationsy voluntarily removed the link, a question arose about whether providing a link to material that infringes on copyright is also copyright infringement. The answer is a legal grey zone that depends heavily on different jurisdictions, and for safety reasons, academics may prefer to just avoid discussing or acknowledging shadow libraries, regardless of which way their personal opinions swing. John Bohannon mentions offhandedly that many of his own papers can be found on Sci-Hub, but he does not give any explicit endorsement or condemnation.

Another reason that may contribute to authors' absence from the discussion is that they don't always have stakes in the situation. As a condition of publishing scholarly work, authors typically transfer their copyright over to the publisher (Anderson, 2018). Furthermore, authors usually don't expect to make money from publishing their work, though royalties tend to be more common for monographs. Since authors are not technically the ones whose rights are being infringed and they are not being affected financially, is there reason for them to care if their work ends up in a shadow library? It would be of interest to examine if attitudes towards shadow libraries significantly differ between scholars in fields that primarily publish journal articles versus those that prioritize monographs. The sheer amount of material uploaded to shadow

libraries may also affect authors' attitudes. Those who would care about their work being infringed on may feel powerless to prevent it.

One method to get at understanding authors' perspectives is extrapolation by looking at the other roles that authors play in the scholarly ecosystem. For example, researchers can be considered as authors; after all, the product of research is typically some form of scholarly publication. When their own work has been supported by research obtained from shadow libraries, authors would presumably be advocates for those resources. On the other hand, researchers have expressed concern over their ability to publish work based on illegally acquired data (Bohannon, 2016). Not everyone is worried about the legal concerns though. In one article cited in this paper, the conflict of interest section discloses "The author uses SciHub regularly in his own research" (Greshake, 2017, p. 7). One can also search Google Scholar for the phrase "thank alexandra elbakyan" and surface dozens of papers that acknowledge their research would not have been possible without the use of Sci-Hub. Additionally, researchers often have a role as professors, and they are invested in their students having access to the sources they need. *Shadow Libraries: Access to Knowledge in Global Higher Education* details the history of photocopying as a way for professors to cheaply provide readings to their students, though unlike shadow libraries, this use is often covered by free use provisions.

However, both of these frameworks involve the use of other people's work and not their own. It's conceivable that an author may be willing to circumvent copyright in their own research while expressing disapproval at others doing the same with their own work. To understand the range of author opinions, social media could be one useful source. An author might post on an academic networking site or on less professional social media like Twitter more freely than they would in a formal, "on the record" setting.

Conclusion

Shadow libraries gather digital copies of copyright protected and paywalled material and make them freely accessible online. Shadow libraries exploded in popularity in response to the demand for affordable access to scholarship created by the serials crisis, with the largest known shadow library being Sci-Hub. Sci-Hub has been the subject of lawsuits from academic publishers, praise from researchers who rely on the site, and a mix of opinions from people who agree with the values of providing knowledge freely but criticize illegal activity. Although scholars have researched shadow library users in detail and librarians have in some cases published opinions on the matter, very little is known about how the authors whose publications end up in shadow libraries feel about their writing being made available there.

Authors may choose not to talk about their opinions publicly due to a variety of factors including but not limited to concerns over legal grey areas, lack of investment after transferring copyright to publishers, and resignation over the scale of infringement. Despite these factors, there is still interest in considering their thoughts and opinions because of the key role that authors play in the scholarly publishing ecosystem. Some opinions may be deduced by examining from the perspective of authors as researchers, while others can be surfaced by looking through debates and other posts on social media. However, a dedicated research study would beget the most useful information. Some proposals for such a study could involve looking at authors published in the most commonly infringed upon disciplines and journals, and interviewing them under the condition of anonymity.

References

Anderson, R. (2018). *Scholarly communication: What everyone needs to know*. Oxford University Press.

Bohannon, J. (2016). Who's downloading pirated papers? Everyone. *Science (American Association for the Advancement of Science)*, 352(6285), 508-512. https://doi.org/10.1126/science.352.6285.508

Greshake, B. (2017). Looking into pandora's box: The content of Sci-Hub and its usage. *F1000Research*, 6:541. https://doi.org/10.12688/f1000research.11366.1

Harrison, R., Nobis, Y., & Oppenheim, C. (2018, November 9). A librarian perspective on Sci-Hub: the true solution to the scholarly communication crisis is in the hands of the academic community, not librarians. *LSE Impact Blog.* https://blogs.lse.ac.uk/impactofsocialsciences/2018/11/09/a-librarian-perspective-on-sci-hub-the-true-solution-to-the-scholarly-communication-crisis-is-in-the-hands-of-the-academic-community-not-librarians/

Karaganis, J. (ed.). (2018). *Shadow libraries: Access to knowledge in global higher education*. MIT Press. https://doi.org/10.7551/mitpress/11339.001.0001

McKenzie, L. (2019, August). *Linking liability*. Inside Higher Ed. https://www.insidehighered.com/news/2019/08/16/legal-questions-raised-over-links-sci-hub

Ottaviani, J. (2016). The post-embargo open access citation advantage: It exists (probably), it's modest (usually), and the rich get richer (of course). *PLoS ONE 11(8)*. https://doi.org/10.1371/journal.pone.0165166

Schiermeier, Q. (2015, December 04). Pirate research-paper sites play hide-and-seek with publishers. *Nature*. https://doi.org/10.1038/nature.2015.18876

Schiermeier, Q. (2017, June 22). US court grants Elsevier millions in damages from Sci-Hub. *Nature*. https://doi.org/10.1038/nature.2017.22196