Carrie Russell and Ed Sanchez

Sci-Hub unmasked

Piracy, information policy, and your library

Academic libraries learned in the summer of 2015 that their expensive Elsevier and Wiley subscriptions were the target of mass copyright infringement by hackers who identified with the Library Genesis Project.[1] Hackers took over university accounts, copied journal content, and set up a searchable repository at a website called Sci-Hub—all to ensure that disadvantaged researchers of the world could get access to content they cannot afford. Because many scholarly publishers charge libraries outrageous prices and have profit margins comparable to the most profitable drug companies, one's first reaction might have been, "Great! It's about time those publishers got their comeuppance."

We may have disdain for scholarly publishers who singularly benefit from a warped scholarly communication system that our university faculty and researchers enable. We may champion the egalitarianism of open access, and even sympathize with modern day hackers who want to share information. However, we also recognize that mass infringement is wrong. At the end of the day, it is likely certain publishers will increase subscription costs to account for this infringement (or at least that will be the argument).

In this article, we will discuss the relative ease with which hackers can access library content and describe concrete steps that libraries can and should take to limit infringement. We will also consider how we place this issue in a broader context. How do we reconcile our belief in equitable access with our own self-interests and our sympathy with the Robin Hood hackers of the world?

The anatomy and consequence of an attack

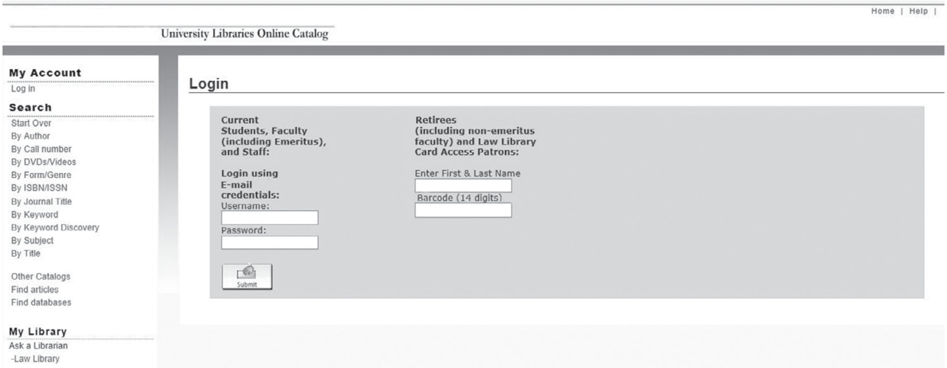
Based in Kazakhstan, Sci-Hub hackers allegedly use compromised user credentials—usernames and passwords—to access proxy servers that manage access to licensed IP-authenticated content from academic institutions. Once access is obtained, the hackers actively gather copyright-protected materials into vast online collections that are then made available via the web to sci-hub.org or libgen.org "customers" around the world.

Sci-Hub takes advantage of an active international market in stolen user credentials, where innocent users give up their passwords to phishing attacks targeting the university community. In one such email attack, the hacker poses as a library service manager by using a combination of two real library staff members' names familiar to faculty. The email draws users to a familiar URL address but, instead of taking them to their own library server, sends them to a secondary page (see Figure 1) with similar branding, though hosted in New Zealand. Input typed into the username and password fields on this page is captured and later used to illegally access licensed content.

Typically, when vendors become aware of a compromise, they make the institution aware of the breach and provide the log files

Carrie Russell is director of the Program on Public Access to Information, ALA's Washington Office, email: crussell@alawash.org, and Ed Sanchez is head of Library and Information Technology at Marquette University, email: edward.sanchez@marquette.edu

so that the institution can begin the process of identifying the compromised user and resolve the problem within a short timeframe specified by the vendors. In most cases, only the IP address of the Sci-Hub user is provided. Additional steps may be necessary to identify the patron ID of the compromised account in order to reset the password. If the investigation is not completed in a timely fashion, vendors may cut off access to their materials. Depending on how quickly an IT team can respond, interrupted access to a popular online resource can become a challenge for library patrons. Further, depending on the level of the access authority of the compromised user, his or her credentials could be used for other potentially nefarious purposes.



Login

Current Students, Faculty (including Emeritus), and Staff:

Login using E-mail credentials:
Username:
Password:

Retirees (including non-emeritus faculty) and Law Library Card Access Patrons:

Enter First & Last Name

Barcode (14 digits)

Figure 1: Sci-Hub mockup of Library Catalog login page. View this article online for more detailed image.

Content provider response to the Sci-Hub incident

In June 2015, one of the major content providers, Elsevier, filed suit in U.S. District Court for the Southern District of New York, naming sci-hub.org, The Library Genesis Project, and Alexandra Elbakyan (believed to reside in Almaty, Kazakhstan) as defendants in a civil action seeking damages and injunctive relief for copyright infringement and for violation of the Computer Fraud and Abuse Act.

In October 2015 the court ruled in favor of Elsevier, agreeing that the defendants fraudulently obtained student or faculty access credentials on university campuses and used those credentials to gain unauthorized access to copyrighted scholarly journals, articles, and books hosted on ScienceDirect.

Another content provider taking action to protect themselves against Sci-Hub is Wiley. In July 2015, Wiley informed customers that Sci-Hub was targeting student and faculty access credentials using methods similar to those mentioned in the Elsevier complaint, and offered guidance on identifying compromised systems and securing them against further attacks. The communication concluded with a list of detection and prevention strategies, including checking for open ports, reviewing firewall logs for communications with any of Sci-Hub's anonymous proxies, and a list of recommended network utilities.[2] Later that month, Wiley went further by announcing the use of a CAPTCHA challenge for download requests exceeding a specified maximum within a 24-hour period.[3]

This is a library issue

Publisher-initiated legal action, strategies for detection, and session download restrictions will have only partial success in limiting the infringement of licensing terms. Moreover, the costs of these measures will eventually be borne by libraries through higher subscription costs.

Librarians can take a leadership role to protect patrons' online credentials and prevent unauthorized intrusion of licensed collections by looking at their authentication processes, monitoring their systems to detect compromised user accounts,[4] and working in conjunction with campus IT departments, publishing and content provider partners, and other agencies and associations toward developing strategies that most effectively address these collective concerns.

Libraries that can link phishing attacks directly to illicit downloads should inform others and take steps to prevent theft of online credentials through strategies like two-factor authentication[5] identity management systems like Shibboleth or CAPTCHA. Ultimately, reliance on password protection in an environment where users are increasingly being fooled into giving them away is counterproductive.

Beyond the local actions, however, there are broad information policy issues that need the attention of the library community.

Why this is also an information policy issue

Sci-Hub and its affiliated sites are not motivated by commercial gain, according to its founder, neuroscientist Alexandra Elbakyan, but rather to remove all barriers in the way of science. Indeed, the ideological and cultural tensions at play in *Elsevier v. Sci-Hub* are not dissimilar to those spurring the debate over open access. While our profession has rightly embraced the open access movement, massive infringement of the kind described is unacceptable. But what can libraries do to increase access to information through open access that might eliminate the desire for people like Sci-Hub to infringe? Or perhaps more broadly stated, what can we do to make the scholarly communication system more equitable/sustainable?

Often the best information policy transformations occur on the ground by the actions of people without the sanction of law or other formal authorization. One can trace the beginnings of the open access movement to 1991 when Paul Ginsparg and a group of physicists who wanted to share their research established the freely accessible pre-print repository arXiv. The movement has gained momentum and precipitated federal agencies' and private foundations' establishment of open access policies requiring unrestricted access and reuse of all peer-reviewed published research funded by those organizations. Although change is slow, it is likely that open access will be key in addressing the problems of unequal access to information. As champions of the movement, libraries can and should lead and take actions such as:

- advocating for open access and changes in information policies through involvement in National Library Legislative Day, and ALA Washington Office and ACRL calls for grassroots work to support or oppose legislation, such as the Fair Access to Science and Technology Research Act (FASTR);[6]
- educating research communities about the Executive Directive on Public Access,[7] open access, author rights, and open licensing through initiatives such as the Scholarly Publishing and Academic Resources Coalition (SPARC);
- actively providing assistance with compliance with funders' open access policies and educating users in the discovery of freely available research materials in open access repositories; and
- reconsidering the allocation of resources for collection development and interlibrary loan in an attempt to steer commercial publishers toward a financially sustainable scholarly communication ecosystem that is beneficial to all stakeholders.

Finally, international efforts such as the World Summit on the Information Society (WSIS) and the World Intellectual Property Organization's (WIPO) Development Agenda[8] have brought the inequities in information access to the public consciousness. The objective of the Development Agenda is to provide technical assistance for countries building their own intellectual property systems, laws, and policies. The Library Copyright Alliance (LCA)—the U.S. library coalition of ALA, ACRL, and Association of Research Libraries (ARL)—has played an active role in these global efforts. Developing countries are eager to listen and learn about U.S. copyright law's exceptions and limitations, particularly fair use. LCA can provide an alternative viewpoint—balanced copyright— to cultural ministers and other government officials who may know copyright only from a maximalist perspective.

Ideally, crossborder sharing of library materials would facilitate access to library materials but this is strongly opposed by rights holders. It is unlikely to be even a consideration as the United States reviews its copyright law in the next few years. But some scholars suggest that our current interlibrary loan exception in section 108 of the copyright law already allows for crossborder sharing as long as the receiving library abides by U.S. copyright law.[9] But will U.S. libraries be willing to push that envelope, when they are stymied by licensing agreements that forbid interlibrary loan? Unfortunately, proposals to change the copyright law to ensure that license agreements do not circumvent library exceptions are unlikely. There is too much money to be made to accommodate sharing, which brings us back to open access.

Conclusion

Taking the security and open access components of the Sci-Hub case forward as discussion items in ALA divisions like ACRL, LITA, and LLAMA may result in best practices and strategies for securing library systems against attack and more insightful discussion of the competing values.

We need to work more closely with organizations like IFLA and promote their Guidelines for International Lending.[10] Even if we merely share our stories with developing nations, they crave the knowledge, and we can provide them with our expertise, effective solutions, and advocacy models, as we are currently doing with public library ebooks.[11]

Working with publishers and information agencies to address security concerns and collaborate in the reduction if not elimination of organized infringement is imperative. We can also raise this issue with ILS vendors with whom we contract to achieve a level of certainty that their authentication systems are up to the challenge of preventing or mitigating these types of cyber-attacks.

Finally, we must promote efforts to fix or replace the current scholarly publishing system by supporting and promoting open access at the local, regional, and national levels. The ubiquity of the Internet will multiply the means for creating and disseminating scholarly research, and this must be harnessed for the good of scholars, publishers, and researchers, but equally important are the questions of information equity around the world.

Notes

1. Library Genesis, 2016, http://gen.lib.rus.ec/.

2. Wiley Customer Information Notice 20150630 (copy on file with author).

3. Email from Wiley, Inc., Wiley Online Library adds CAPTCHA to prevent systematic article downloads, July 9, 2015, http://app.news.wiley.com/e/es?s=1133198723&e=227111&elq=2f601287796240e9b7a884ee213e4538

4. James Padgett and Jonathan Hooper, "SierraDNA—Demonstrating the Usefulness of Direct ILS Database Access," *Code{4}lib Journal* 30 (10/15/2015). http://journal.code4lib.org/articles/10924.

5. Thomas Sui, "On the move with two-factor authentication," EDUCAUSEreview, (October 2015), http://er.educause.edu/blogs/2015/10/on-the-move-with-two-factor-authentication.

6. Alliance for Taxpayer Access, "Fair Access to Science and Technology Research Act," 2015, www.taxpayeraccess.org/action/FASTR/index.shtml.

7. Executive Directive on Public Access –SPARC, "About the Directive," 2016, http://sparcopen.org/our-work/2013-executive-directive/.

8. WIPO Development Agenda: http://www.wipo.int/ip-development/en/agenda/.

9. Brandon Butler and Kevin Smith, "White paper: U.S. law and international library loan," *Research Library Issues* 275 (June 2011).

10. International Resource Sharing and Document Delivery: Principles and Guidelines for Procedure, 2009 Revision, www.ifla.org/publications/international-resource-sharing-and-document-delivery-principles-and-guidelines-for-proc.

11. Julia Brings, Committee on Copyright and other Legal Matters (CLM), "IFLA Statement on Public Lending Right Updated to Include eBooks," (February 1, 2016) http://www.ifla.org/node/10205.