Unpaywall finds free versions of paywalled papers

New tool joins a growing collection of software for accessing fee-for-view scientific literature.

Dalmeet Singh Chawla

04 April 2017

An online widget that trawls the Internet searching for free-to-read versions of paywalled papers has been installed more than 10,000 times since its prerelease debut on 10 March, its inventors say.

Officially launched on 4 April, Unpaywall is a free web-browser extension that hunts for papers in more than 5,300 repositories worldwide, including preprint servers and institutional databases.

Once installed in Google Chrome or Mozilla Firefox, Unpaywall brings up a green or grey tab on the side of the screen when a user hits a paywalled paper. A green 'unlock' sign means that a free version of the paper is available elsewhere, whereas a grey 'lock' icon means that the tool could not locate a free version.

For researchers and librarians, Unpaywall and a growing collection of similar tools promise to simplify the process of finding freely accessible (and legal) copies of research articles.

Holly Bik, a marine biologist at the University of California, Riverside, says Unpaywall is especially useful when she is off campus. "I can quickly access and download PDFs of journal articles without having deal with the hassle [of] logging into a VPN connection to access my university journal subscriptions," she told *Nature*. And author Nicola Twilley, a contributing writer to *The New Yorker* who is based in Brooklyn, New York, took to Twitter on 23 March to point out that Unpaywall found her a free version of a study half an hour after she installed it. "I am constantly looking for ways to get hold of academic papers as source material," she wrote in an e-mail.

Rick Anderson, associate dean for collections and scholarly communication at the University of Utah in Salt Lake City, says that Unpaywall could alter the calculus that librarians use when choosing subscriptions.

"For years, libraries have had little expectation that the [open access] deposit of subscription articles would displace subscription access, mainly because locating those articles and providing access to them would be labor-intensive," Anderson says. "By making that process so much easier, Unpaywall could really change the game for us."

Bots galore

Of course, readers have had the ability to find article PDFs online for years, thanks to Google and Google Scholar. But these tools do not always retrieve free articles, says Jason Priem, Unpaywall's co-inventor. That's where Unpaywall and other services come in.

At its heart, Unpaywall provides a simple interface into a database of digital object identifiers (DOIs) — some 86 million in all. In October 2016, Priem and co-inventor Heather Piwowar, who together founded the not-for-profit company Impactstory in Vancouver, Canada, launched oaDOI, an 'application programming interface' (API), or set of programmatic instructions, that allows researchers to search that database for papers on the basis of their DOI. oaDOI is used more than a million times per week, according to Priem; it successfully finds an open-access version of a paper around 30% of the time. Unpaywall basically queries that same database automatically, flagging articles for which a free version is available.

oaDOI, in turn, draws inspiration from DOAI (Digital Open Access Identifier), says Priem. DOAI is described on its homepage as an "alternate DOI resolver" that redirects requests to the free version of an article, if available, and otherwise to the paywalled paper itself.

According to Piwowar, both oaDOI and Unpaywall include a nimble title-matching feature, which uses an algorithm to compare titles, forgiving spacing and punctuation differences. The idea is to overcome the fact that titles change in different versions of manuscripts, she explains.

More than 600 libraries worldwide are using oaDOI, Priem says. The National Library of Sweden, for instance, uses oaDOI in their SwePub analysis system, which integrates the metadata of research published at all of Sweden's research institutions.

"By using oaDOI we have been able to increase our statistics on green open access by 38% and gold open access by 27%," says Jürgen Kerstna, formerly a system developer at the National Library of Sweden. (Green open-access, also called self-archiving, refers to authors depositing their papers in online repositories; gold open-access is when journal articles are made available by the publisher.)

Other tools for mining the open-source literature have been or are being developed. These include the bibliographic databases oaFindr and oaFindr+; the Open Access Button, a web-browser extension that searches for open-access papers and allows users to e-mail authors when it cannot find one; and 'pirate' services such as Sci-Hub.

Joseph McArthur, who is assistant director of the Right to Research Coalition policy advocacy group in London, says Open Access Button, which he co-founded, has adopted oaDOI's freely available API as well as additional aggregators to expand its reach. "We're hoping our projects can work in tandem to provide the community with the best tools possible for legally getting around paywalls," he says.

Another in-development tool, OAbot (Open Access Bot), will search the web for free-to-read versions of papers mentioned in references on Wikipedia pages, says Jake Orlowitz, head of the Wikipedia Library at the Wikimedia Foundation in the San Francisco Bay Area. Once fully launched, references on English Wikipedia pages will be marked with a green unlock icon when a free version of the paper is available. In time, Orlowitz notes, the team aims to use multiple icons to better reflect a paper's availability — namely, whether it is paywalled, not paywalled but requires registration, free to read but not free to reuse, or free to read as well as reuse.

If publishers move toward an open-access model, Unpaywall and its ilk would no longer be needed. Priem, for one, would welcome that change, he says: "I'll be pulling the plug with a great big smile on my face."