UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

| | | |
|---|---|---|
| HACHETTE BOOK GROUP, INC., HARPERCOLLINS PUBLISHERS LLC, JOHN WILEY & SONS, INC., and PENGUIN RANDOM HOUSE LLC, | : | 20 Civ. _____ |
| | : | ECF Case |
| Plaintiffs, | : | |
| | : | **COMPLAINT** |
| -against- | : | TRIAL BY JURY DEMANDED |
| | : | |
| INTERNET ARCHIVE and DOES 1 through 5, inclusive, | : | |
| | : | |
| Defendants. | | |

Plaintiffs Hachette Book Group, Inc. ("Hachette"), HarperCollins Publishers LLC ("HarperCollins"), John Wiley & Sons, Inc. ("Wiley"), and Penguin Random House LLC ("Penguin Random House"), by and through their attorneys Davis Wright Tremaine LLP and Oppenheim + Zebrak, LLP, for their Complaint, hereby allege against Defendant Internet Archive ("IA" or "Defendant") and Does 1 through 5 as follows:

## NATURE OF THE ACTION

1. Plaintiffs Hachette, HarperCollins, Penguin Random House, and Wiley (collectively, "Plaintiffs" or "Publishers") bring this copyright infringement action against IA in connection with website operations it markets to the public as "Open Library" and/or "National Emergency Library." Plaintiffs are four of the world's preeminent publishing houses. Collectively, they publish some of the most successful and leading authors in the world, investing in a wide range of fiction and nonfiction books for the benefit of readers everywhere. All of the Plaintiffs are member companies of the Association of American Publishers, the mission of which is to be the voice of American publishing on matters of law and public policy.

2. Defendant IA is engaged in willful mass copyright infringement. Without any license or any payment to authors or publishers, IA scans print books, uploads these illegally scanned books to its servers, and distributes verbatim digital copies of the books *in whole* via public-facing websites. With just a few clicks, any Internet-connected user can download *complete* digital copies of in-copyright books from Defendant.

3. The scale of IA's scheme is astonishing: At its "Open Library," located at www.openlibrary.org and www.archive.org (together, the "Website"), IA currently distributes digital scanned copies of over *1.3 million* books. And its stated goal is to do so for millions more, essentially distributing free digital copies of every book ever written. Despite the "Open Library" moniker, IA's actions grossly exceed legitimate library services, do violence to the Copyright Act, and constitute willful digital piracy on an industrial scale. Consistent with the deplorable nature of piracy, IA's infringement is intentional and systematic: it produces mirror-image copies of millions of unaltered in-copyright works for which it has no rights and distributes them in their entirety for reading purposes to the public for free, including voluminous numbers of books that are currently commercially available.

4. Books have long been essential to our society. Fiction and non-fiction alike, they transport us to new worlds, broaden our horizons, provide us with perspective, reflect the ever-growing knowledge of humanity in every field, spark our imaginations and deepen our understanding of the world. Yet, books are not self-generating. They are the product of training and study, talent and grit, perseverance and creativity, investment and risk, and untold hours of work.

5. The publishing ecosystem not only depends upon copyright law, it is historically intertwined with the founding of the United States. In 1787, the Framers adopted the Copyright

Clause of the Constitution, explicitly authorizing Congress "[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries." U.S. Const., Art. I, §8, cl. 8. In 1790, the First Congress enacted the first Copyright Act, focused on incentivizing both the creation *and legal dissemination* of books, maps, and charts. Congresses ever since have carefully balanced copyright amendments to advance the public good and for more than 200 years have prescribed to authors a suite of enforceable exclusive rights to their writings—which publishers, in turn, encourage, invest in, license, and distribute to readers through bookstores, libraries, and a multitude of e-commerce platforms. In this process of publishing books that educate, entertain, and inspire the public, publishers rely not only on the exclusive rights that are their lifeblood, but on the expectation that Congress has carefully considered and appropriately tailored any limitations and exceptions to said rights.

6. IA not only acts entirely outside any legal framework, it does so flagrantly and fraudulently. And it proceeds despite actual notice that its actions constitute infringement. For the avoidance of doubt, this lawsuit is *not* about the occasional transmission of a title under appropriately limited circumstances, nor about anything permissioned or in the public domain. On the contrary, it is about IA's purposeful collection of truckloads of in-copyright books to scan, reproduce, and then distribute digital bootleg versions online. IA's Website includes books of every stripe—from bestsellers to scholarly monographs, from entertaining thrillers and romances to literary fiction, from self-help books to biographies, from children's books to adult books. IA often suggests that the Website is limited to twentieth-century books, but this is neither accurate nor a defense. IA scans, uploads, and distributes huge numbers of in-copyright books published in both the twentieth and twenty-first centuries, including many books

published within just the past few years. IA's unauthorized copying and distribution of Plaintiffs' works include titles that the Publishers are currently selling commercially and currently providing to libraries in ebook form, making Defendant's business a direct substitute for established markets. Free is an insurmountable competitor.

7. Publishers have long supported public libraries, recognizing the significant benefits to the public of ready access to books and other publications. This partnership turns upon a well-developed and longstanding library market, through which public libraries buy print books and license ebooks (or agree to terms of sale for ebooks) from publishers, usually via book wholesalers or library ebook aggregators. IA's activities are nothing like those of public libraries, but rather the kind of quintessential infringement that the Copyright Act directly prohibits. Moreover, while Defendant promotes its non-profit status, it is in fact a highly commercial enterprise with millions of dollars of annual revenues, including financial schemes that provide funding for IA's infringing activities. By branding itself with the name "Open Library," it thus badly misleads the public and boldly misappropriates the goodwill that libraries enjoy and have legitimately earned.

8. IA defends its willful mass infringement by asserting an invented theory called "Controlled Digital Lending" ("CDL")—the rules of which have been concocted from whole cloth and continue to get worse. For example, at first, under this theory IA claimed to limit the number of scanned copies of a title available for free download at any one time to the number of print books of that title in its collection—though no provision under copyright law offers a colorable defense to the systematic copying and distribution of digital book files simply because the actor collects corresponding physical copies. Then, in the face of the COVID-19 pandemic, IA opportunistically seized upon the global health crisis to further enlarge its cause, announcing

with great fanfare that it would remove these already deficient limitations that were purportedly in place. Today, IA offers an enormous universe of scanned books to an unlimited number of individuals simultaneously in its "National Emergency Library." IA's blatant, willful infringement is all the more egregious for its timing, which comes at the very moment that many authors, publishers, and independent bookstores, not to mention libraries, are both struggling to survive amidst economic uncertainty and planning deliberatively for future, changing markets.

9. Under whatever guise IA attempts to frame its massive infringement—whether adopting the invented CDL theory or filling the self-appointed role as "National Emergency Library"—its actions find no support in the Copyright Act. IA's defenses of its actions—both before and after the onset of the COVID-19 crisis—are baseless. First, while IA claims to serve an educational purpose, education has long been a primary mission and market of publishers. It is authors and publishers who create the books of scholarship and literature for educators, students, and other readers; IA creates nothing. IA plays no role in the hard work of researching, writing, or publishing the works or, for that matter, in creating or sustaining the overall publishing ecosystem and its distinct partnerships and markets. Nor does IA contribute to the underlying scholarship through commentary or criticism. Moreover, IA's massive book digitization business has no new purpose that is fundamentally different than that of the Publishers: both distribute entire books for reading. In short, Defendant merely exploits the investments that publishers have made in their books, and it does so through a business model that is designed to free-ride on the work of others. Defendant pays for none of the expenses that go into publishing a book and is nothing more than a mass copier and distributor of bootleg works. In so doing, IA undermines the balance and promise of copyright law by usurping the

DOCKET ALARM

# Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

## Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

## Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

## Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

## API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

**LAW FIRMS**
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

**FINANCIAL INSTITUTIONS**
Litigation and bankruptcy checks for companies and debtors.

**E-DISCOVERY AND LEGAL VENDORS**
Sync your system to PACER to automate legal marketing.