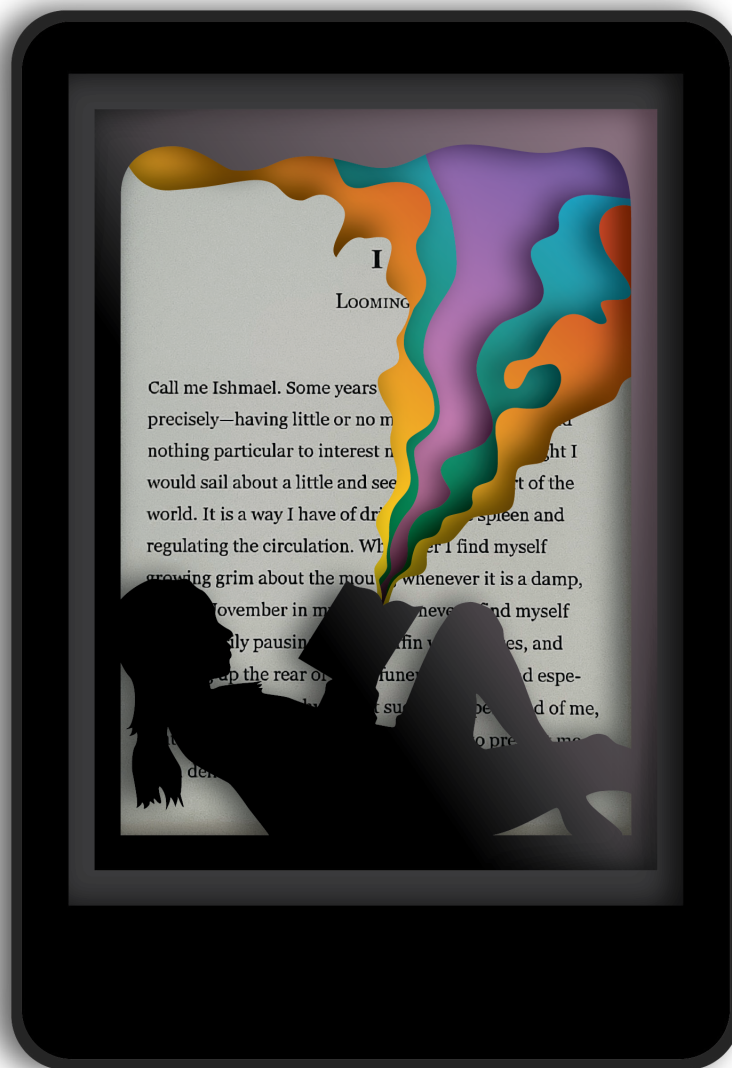


# The Anti-Ownership Ebook Economy

How Publishers and Platforms Have  
Reshaped the Way We Read in the Digital Age



## Engelberg Center on Innovation Law & Policy

The Engelberg Center on Innovation Law & Policy at New York University School of Law provides a unique environment where scholars can examine the key drivers of innovation as well as the law and policy that best support innovation. By fostering interdisciplinary and collaborative research on innovation law and policy, the Engelberg Center attracts legal scholars and practitioners, technologists, economists, social scientists, physical scientists, historians, innovators, and industry experts who study the incentives that motivate innovators, how those incentives vary among creative endeavors, and the laws and policies that help or hinder them.

## Acknowledgements

Thank you Claire Ewing-Nelson, Leslie Peters, all of the stakeholders who generously made time to be interviewed for this project, and everyone who provided feedback, insights, and suggestions on drafts.

## Authors

### SARAH LAMDAN

Sarah Lamdan is a Professor of Law at CUNY School of Law who specializes in information law. Her research and advocacy span the spectrum from public information access to personal data privacy. Sarah Lamdan is also a law librarian with a master's degree in library and information science. When she's not teaching, she works with organizations like SPARC, Library Futures, Library Freedom Project, and other open access and data privacy advocacy groups.

### JASON M. SCHULTZ

Jason M. Schultz is a Professor of Clinical Law, Director of NYU's Technology Law & Policy Clinic, and Co-Director of the Engelberg Center on Innovation Law & Policy. His clinical projects, research, and writing primarily focus on practical frameworks and policy options to help traditional areas of law such as intellectual property, privacy, consumer protection, and civil rights adapt in light of new technologies and the challenges they pose. His most recent work focuses on the social and legal implications of machine learning, artificial intelligence, and the Internet of Things.

### MICHAEL WEINBERG

Michael Weinberg is the Executive Director of the Engelberg Center on Innovation Law and Policy at the NYU School of Law. His work focuses on open source, open access, and the intersection of society and emerging technologies.

### CLAIRE WOODCOCK

Claire Woodcock is an independent journalist and recent graduate of the University of Colorado Boulder's Media & Public Engagement program. Her work has appeared in Slate, Vice, NPR, and other publications.

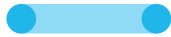


# Contents

<b>Introduction</b>	<b>4</b>
<b>The Emergence of the Publisher-Platform Partnership</b>	<b>8</b>
<b>From Physical Book Sales to Ebook Licensing – An Opportunity for Publishers to Achieve Long-Held Goals</b>	<b>11</b>
<b>Ebooks Reduce the Friction in Library Lending, Raising Concerns for Publishers</b>	<b>24</b>
<b>The Publisher-Platform Partnership Shapes the Ebook Market</b>	<b>33</b>
<b>Monetizing Data and Tracking Users Provide Additional Incentives for Publishers and Platforms to Retain Control</b>	<b>44</b>
<b>Conclusions</b>	<b>51</b>



# Introduction



Have you ever noticed that you can't really "buy" an ebook? Sure, when you click that "Buy Now" button on your ereader, tablet, or phone, it feels like a complete, seamless transaction. But the minute you try to treat your ebook like a physical book – say by sharing it with a friend, selling it to someone else, donating it to a school library, or sometimes even reading it offline, reality sets in. You can't do any of those things.

With most ebooks, even if you think you "own" them, the publisher or platform you bought them from will say otherwise. Publishers and platforms insist that you only buy a license to access the books, not the rights to do anything else with them. And because platforms like Amazon and Apple control most of the technology we use to read ebooks, their opinion often dictates the reality of the ebook ecosystem. Beyond controlling books, these platforms can also do several things no physical bookseller has ever had the power to do. They can track your reading habits, stop you from reselling or lending a book, change the book's content, and delete it from your digital library altogether – even after you've bought it. This doesn't happen in the print book market, where you can still feel confident that when you buy a book, it's yours to share, sell, or simply read without it being tracked or censored.

Something happened when we shifted to digital formats that created a loss of rights for readers. Pulling back the curtain on the evolution of ebooks offers some clarity to how the shift to digital left ownership behind in the analog world.

While most publishers still sell physical books, when it comes to ebooks, the vast majority appear to have made a collective decision to shift to offering only limited licenses. Some of the reasons for this shift are economic, some legal, some technological, and others psychological – a belief that limiting or eliminating digital ownership of books will raise publisher revenues, forstall free copies leaking onto unauthorized websites, and allow publishers and platforms unprecedented control and tracking of the behaviors of readers, as well as universities and libraries that provide ebooks. Whether these beliefs map to reality, however, is hotly contested.

**Just as platforms control our tweets, our updates, and the images that we upload, platforms can also control the books we buy, keeping tabs on how, when, and where we use them, and at times, modifying or even deleting their content at will.**



Economically, platforms and publishers also believe licensing gives them more control over how ebooks generate revenues. For example, they currently use licensing to charge libraries for the right to lend the book to patrons, or tying ebooks to platforms that monitor user behavior, creating a new revenue stream based on selling user data.

Legally, the shift from selling to licensing attempts to circumvent centuries of law that have limited publisher control over post-sale uses of books. This law is a copyright doctrine called “exhaustion” or “first sale.” The idea behind first sale is that publishers were always entitled to make money from the first time they sold a book, but after that, the sold copy is beyond their control. The new owner can decide to resell, lend, or use the book in any manner they see fit. Licensing attempts to keep that control with the copyright owner, forever.

Technologically, publishers have turned to companies like Amazon, Apple, and OverDrive to distribute and control their ebooks. These publisher-platform partnerships presented new opportunities for publishers to remain involved in the post-sale life of a book. Just as platforms control our tweets, our updates, and the images that we upload, platforms can also control the books we buy, keeping tabs on how, when, and where we use them, and at times, modifying or even deleting their content at will. That same control also allows platforms to monitor when and how you read, which is key to new revenue streams based on selling user data.

Finally, psychologically, publisher-platform partnerships have reinforced a belief long held by publishers that with greater legal and technological control, they will reap greater financial rewards. Publishers believe these rewards flow from forcing purchasers to pay for additional uses (such as when libraries want to lend out ebooks to patrons), and by limiting how books are shared and distributed outside of their control. While it is unclear that any of these beliefs have manifested into reality, the psychology has taken hold, in part because ebook platforms have now locked publishers into this partnership in ways that make it extremely difficult to disentangle.

This report explains that, while there is nothing new about publishers’ desire to seek novel ways to increase revenues, along with control and surveillance of readers, the new publisher-platform partnership creates a mechanism to align the ebook market with those goals. That new market alignment raises questions about whether these shifts are the best option for readers and institutional book buyers, particularly libraries. It also raises questions about how the newest players in the market – ebook distribution platforms – shape things to align with their own interests.

In order to fully understand the dynamics at play, we interviewed over 30 stakeholders that fill various essential roles in the ebook marketplace, from publishers to platform CEOs to literary agents, librarians, and lawyers.<sup>1</sup> We discussed the priorities, concerns, and constraints that help shape their participation in the ebook marketplace. Our goal was to understand and document how this world looks through their eyes, and synthesize those views into broader conclusions. In the following sections, we will deconstruct the shifts that have made it difficult to actually own an ebook collection, as well as the challenge that lies ahead if we wish to resurrect ownership as a consumer choice for ebooks.<sup>2</sup>

Our study leads us to several key conclusions:

- By turning to platforms as the primary technical means for conveying ebooks, publishers have introduced a third major player into the ebook supply chain: ebook platform companies. Together with publishers, platforms have restricted the ebook market to one composed primarily of licensing instead of sales.
- The platform companies have motives and goals that are independent of those of publishers or purchasers (including institutional buyers such as libraries and schools). Rather than looking to profit from individual sales, like a bookstore does, platforms compete to collect and control the most aggregate content and consumer data. This enables what are now widely known as “surveillance capitalism” revenue models, from data brokering to personalized ad targeting to the use of content lock-in subscription models.<sup>3</sup> These platforms’ goals are sometimes at odds with the interests of libraries and readers.
- The introduction of platforms, and especially publisher-platform partnerships, has created new forms of legal and technological lock-in on the publisher side, with dependencies on platform infrastructure posing serious barriers to publishers independently selling ebooks directly to consumers. Platforms have few incentives to support direct sales models that do not require licensing, as those models do not easily support tracking user behavior.

---

1 Many sources expressed fear of retaliation from their employers or the industry at large for participating in this study; therefore, steps have been taken to de-identify research study participants.

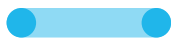
2 There is significant confusion over terminology used to describe the ebook market. Ebook platforms, and stakeholders involved in all aspects of ebook production and sales, use the language of “buying” and “selling” to describe almost all ebook marketplace transactions. This is the same language they use for physical books. However, this shorthand obscures the very real differences among buying a physical book, licensing an ebook, and buying an ebook. This paper tends to use more precise descriptions of various exchanges throughout.

3 See Sarah Lamdan, *Data Cartels* (2022); Shoshana Zuboff, *The Age of Surveillance Capitalism* (2019); Press release, FEDERAL TRADE COMMISSION, <https://www.ftc.gov/news-events/news/press-releases/2021/10/ftc-ramp-enforcement-against-illegal-dark-patterns-trick-or-trap-consumers-subscriptions> (last visited Mar. 16, 2023).

- The structure of the ebook marketplace has introduced new stressors into both the publishing and library professions. Publishers and libraries feel they are facing existential crises/collapse, and their fears are pushing them into diametrically opposed viewpoints. Publishers feel pressured to protect and paywall their content, while libraries feel pressure to maintain relevant collections that are easily accessible via digital networks. Both libraries and publishers feel dependent on the ebook platform companies to provide the ebooks that readers demand, allowing the platform economy (which is already dominated by only a few large companies) to have even more power over the ebook marketplace.
- Because of the predominance of the publisher-platform licensing model for the ebook marketplace, important questions exist as to the impact, if any, that digital library lending of books has on that market. For example, while some evidence exists that the availability of second-hand physical books via libraries and used bookstores might compete with direct publisher book sales, it is less clear that the digital loan of a single title by a library competes with platform ebook subscriptions and locked-in book purchases. Moreover, given that publisher-platform partnerships profit from surveillance of book buyers, consumers who choose more privacy-friendly library loans may represent an entirely distinct market that places significant value on data protection.
- While access to user data generated by platform surveillance of readers is a potential benefit to publishers, in practice publishers do not fully exploit (and may not have full access to) that information.



# The Emergence of the Publisher-Platform Partnership



Before getting into the details of how the ebook marketplace has evolved, it is worth considering the major market players and how they relate to one another.<sup>4</sup> This is key to understanding why the publisher-platform partnership has been so effective in structuring the ebook market in a way that is different from the physical book market.

## The Physical Book Market

One simplified way to imagine the physical book market is as having two poles. One pole is publishers. The other is purchasers.

The publishers work with authors to bring individual titles to market. It is the role of the publishers to manage the economic life of the book. Most authors hand over control of the economic life of the book – marketing, design, manufacturing, distribution – to the publishers.

---

<sup>4</sup> This report is primarily focused on how larger publishers have approached the ebook market. Smaller publishers have experimented with a much-broader range of ebook distribution models, including perpetual licenses (often described as “One Copy/One User” or “OC/OU”). However, they are also a relatively small portion of the market when compared to the combined size of the larger publishers.

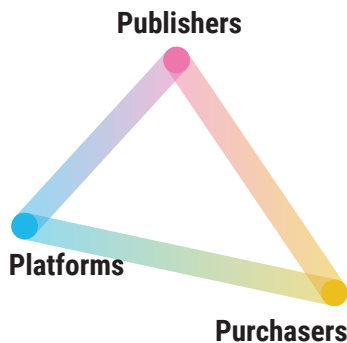




**Figure 1. The physical book market.** One simplified way to imagine the physical book market is as having two poles. One pole is publishers. The other is purchasers.



**Figure 2. Bookstores serve as a bridge between publishers and purchasers.** Bookstores order books from publishers, market them to purchasers, and ultimately deliver them into the hands of purchasers.



**Figure 3. The ebook market and the publisher-platform partnership.** Instead of the bipolar publisher-purchaser physical book market, the ebooks market involves publishers, purchasers, and platforms. These nodes are in constant dialogue with each other, well past the point of purchase.

The purchasers take a book out of the publisher’s control and into the world. The purchaser may be an institutional purchaser, like a library, or an individual reader who wants to own the book themselves. The purchaser may also provide others access to the books, such as public school students or library patrons. Whatever form they take, once the purchaser has purchased the book, the book leaves the control of the publisher and becomes the property of the purchaser.

Of course, bookstores (and equivalent intermediaries in the institutional context) are also an important part of this equation. They serve as a bridge between publishers and purchasers. Bookstores order books from publishers, market them to purchasers, and ultimately deliver them into the hands of purchasers.

However, bookstores have only a transitory moment of control over the books. When they order books, they either sell them to purchasers or return them to publishers. The bookstores have no direct control over how publishers decide to create individual titles, nor do they have the ability to control how purchasers use the books, once they are purchased. In fact, all of the flows in the physical book market are essentially one way – from publishers, to bookstores, to purchasers.

## The Ebook Market and the Publisher-Platform Partnership

The ebook market is different in some key ways. One important difference is that the flows in the ebook market do not move in just one direction. Swapping ebook platforms for bookstores creates a long-term connection among everyone involved in the transaction. Unlike distributors and bookstores, ebook platforms have an ongoing relationship with ebooks that extends well beyond the point of purchase. Not only do ebook platforms market and sell the books – they are where purchasers go to manage and read them.

This ongoing relationship with the ebook is fundamentally different from the temporary transactional nature bookstores play in the physical book market. Instead of the bipolar publisher-purchaser physical book market, the ebooks market involves publishers, purchasers, and platforms. These nodes are in constant dialogue with each other, well past the point of purchase.

By maintaining contact with the ebook and the purchaser in perpetuity, platforms have the opportunity to generate new types of data about purchasers and their habits. Publishers and platforms team up to exercise levels of control over purchasers and copies of ebooks in ways that are impossible in the physical book market.

Like all partnerships, the publisher-platform partnership encompasses distinct incentives that are often – but not always – aligned.<sup>5</sup> When publishers and platforms coordinate, they can shape the market in their favor by limiting traditional rights of purchasers.

Publishers and platforms share incentives to increase control over how ebooks are used, especially at the expense of the purchasers. Both can also benefit from accessing reader data, although, as discussed in more detail below, the nature of the benefits differs somewhat. Publishers and platforms also have an interest in preventing purchaser behavior that they disapprove of. Sometimes, that behavior is illegal behavior, such as large-scale pirating of ebooks. Other times, that behavior is simply behavior that the publishers and platforms see as contrary to their own interests, such as lending an ebook to a friend.

There are also times when the publishers' and platforms' interests do not align. Platforms have clear incentives to lock purchasers into their platform, and monetizing user data is core to their business models. Publishers have incentives to sell more books, which could potentially benefit from a diversity of platforms competing with each other for different types of purchasers. Publishers may also be less effective than platforms at fully exploiting user data, and less interested in doing so.

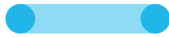
While these differences are important, the aligned interests of the publisher-platform partnership have been powerful enough to create an ebook market that is very different from the physical book market. This paper focuses on a number of dynamics that flow from the partnership, and changes that those dynamics create.



---

5 There are cases in which publishers and platforms are the same entity, with Amazon being the most prominent example. In those cases, while the publishing role and platform role operate at least somewhat distinctly, at least in theory, the interests involved are directly unified within the interests of the larger corporate entity.

# From Physical Book Sales to Ebook Licensing – An Opportunity for Publishers to Achieve Long-Held Goals



The markets for ebooks and physical books are very different. Some of those differences are obvious – physical books are easy to measure and identify as individual objects, while ebooks are more fluid and ephemeral as data that can be stored, transmitted, and accessed in myriad ways. Other differences emerge only when readers attempt to treat ebooks like physical books, and in doing so run into legal and technological restrictions.

As we will explain below, physical books are governed by rules of property and ownership that can be traced back centuries. These rules give people who own physical books clear rights over how they can use those books. However, the rules that govern ebooks are much less clear.

## From Owning to ‘Purchasing’ (With Licenses and Tracking)

For centuries, the law has always balanced the right of an author or publisher to control their intellectual property with the right of a book purchaser to control the copy they buy. Every sale of a physical book involves an exchange. Purchasers pay money to the publisher. In return, publishers give up their ownership over that copy of the book being sold. The purchaser owns their copy of the book and can treat it as their private property, not worrying what the publisher might want.

Over the years, publishers have made many attempts to avoid this exchange, controlling both the purchase price and what purchasers do with the books after they are sold. For example, in the early 1900s, publishers tried to control resale prices on the books people bought from retailers by stamping mandatory resale prices on a book's front page.<sup>6</sup> (That attempt was rejected by the US Supreme Court).<sup>7</sup> Publishers also tried to limit where people could resell books they bought, in one case claiming that a book sold in Thailand couldn't be resold in the US.<sup>8</sup> (That attempt was also rejected by the US Supreme Court, in 2013).<sup>9</sup> These attempts failed because the publisher's copyright does not give them absolute control of a book in perpetuity; the copyright system is a balance between publishers and purchasers.<sup>10</sup> If publishers want the benefits of the copyright system, they also have to accept the limits it imposes on their power.

Publishers do not get to tell purchasers what they can do with physical books, because of a longstanding principle in copyright law called "exhaustion," also known as "the First Sale Doctrine."<sup>11</sup> This rule says that as soon as a copyright holder parts with a copy of their work – whether through a sale, a donation, or even by throwing copies in a trash bin – they no longer have any claim of control over that copy. Instead, whomever comes into lawful possession of that copy now has all of the "incidents of ownership" that come with any other form of personal property. This rule is not limited to physical objects that are protected by copyright. It protects anyone who purchases any sort of physical object. In the same way that a car company cannot shut down independent used car lots and toaster companies cannot police the prices of used toasters at thrift stores, publishers cannot claim control over the copies of books we lawfully own.

---

6 *See Bobbs-Merrill Co. v. Straus*, 210 U.S. 339, 341 (1908) (describing the Bobbs-Merrill Company's practice of including a notice on copies of *The Castaway* stating that the retail price for the book was one dollar net, and that sale at a lesser price would constitute infringement of the company's copyright).

7 *Id.* at 351 ("the copyright statutes ... do not create the right to impose, by notice, such as disclosed in this case, a limitation at which the book shall be sold at retail by future purchasers, with whom there is no privity of contract").

8 *Kirtsaeng v. John Wiley & Sons, Inc.*, 568 U.S. 519, 527 (2013) (federal copyright infringement lawsuit brought by Wiley against individual defendant who purchased English-language textbooks in Thailand and resold them in the United States at a higher price).

9 *Id.* at 525 (holding that "the 'first sale' doctrine applies to copies of a copyright work lawfully made abroad").

10 *Cf. UMG Recordings, Inc. v. Augusto*, 628 F.3d 1175 (9th Cir. 2011) (rejecting attempt by a music company to claim the ability to control resale of promo CDs, even though no records were ever kept of where they were sent and no attempts were ever made to facilitate their return).

11 *See generally* Aaron Perzanowski and Jason Schultz, *The End of Ownership* (2016).

**Copyright law provides incentives for publishers to help produce books, periodicals, and magazines. It also supports institutions such as libraries and archives that work to preserve those works after they are sold and ensure long-lasting access.**



This rule exists for many reasons – respect for personal property, market efficiency, privacy, and preservation chief among them. Copyright law provides incentives for publishers to help produce books, periodicals, and magazines. It also supports institutions such as libraries and archives that work to preserve those works after they are sold and ensure long-lasting access. For physical books, all of this happens seamlessly every day because of the First Sale Doctrine.

The advent of digital books has reinvigorated this long-running battle over post-exchange uses, with publishers claiming that they have the right to charge money for an ebook and decide what happens to that digital copy after money has changed hands.<sup>12</sup>

## Publishers’ Belief That Secondary Markets and Library Lending Undermine Profits

For decades, publishers have expressed displeasure over the ways the First Sale Doctrine has prevented them from exercising control over books once they were sold to customers. These critiques often target the secondary market, where used books are sold at prices, times, and places beyond publishers’ control by sellers with no obligations to report those sales back to the original publishers. After several failed attempts to challenge the First Sale Doctrine in the context of the physical book market, many of the publishers’ concerns seem now to be manifesting within the rules of ebook distribution platforms and licenses.<sup>13</sup>

Economic studies paint a complex picture of the interplay among used book markets, new book sales, publisher profits, and overall consumer welfare for physical books.<sup>14</sup> In some cases, the used book market may increase publisher profits because people are willing to pay a higher price for a new book they know they can sell later. In other cases, a lower-priced used copy of a book could steer

12 Although this report focuses primarily on how the shift to digital has impacted rights connected to the First Sale Doctrine, ebook licenses also attempt to limit other rights, including fair use (17 U.S.C. § 107) and special rights afforded to libraries and archives (17 U.S.C. § 108).

13 Although publisher wariness of the used book market is often the most-visible manifestation of the desire to control books after they are purchased, publishers have also expressed concerns about the simple act of lending books between friends. This concern is discussed in more detail below.

14 See, e.g., Anindya Ghose, Rahul Telang, and Ramayya Krishnan, Effect of Electronic Secondary Markets on the Supply Chain, 22 *J. of Mgmt. Info. Sys.* 91 (2005); Anindya Ghose, Michael D. Smith, and Rahul Telang, Internet Exchanges for Used Books: An Empirical Analysis of Product Cannibalization and Welfare Impact, 17 *Info. Sys. Rsch.* 3 (2006).

**Regardless of the actual impact that a robust secondary market may have on an individual publisher's sales and profits, for well over a century publishers have demonstrated that they believe the secondary market acts to undercut prices, sales, and profits in the primary book market.**



a purchaser away from paying full price for a new copy. However, even in the case of consumer preferences for cheaper used copies, or borrowing books from libraries instead of purchasing them, it is unclear how often used book purchases or library loans are actual substitutes for new-book buying. Participants who work with libraries at Big Five publishing houses did recognize that libraries are sites of discovery that lead to book sales. There are also other social benefits of secondary markets, such as privacy and preservation, that can impact the analysis.

Regardless of the actual impact that a robust secondary market may have on an individual publisher's sales and profits, for well over a century publishers have demonstrated that they believe the secondary market acts to undercut prices, sales, and profits in the primary book market. This belief that the secondary market is harming their business has driven publishers' behavior since at least the start of the 20th century. That publisher wariness of the uncontrolled use of books is key to understanding publisher motivations as the ebook market has evolved.

*Bobbs-Merrill Company v. Straus*, the 1908 Supreme Court case that established the First Sale Doctrine in United States common law, flowed directly from a publisher's attempt to control the minimum price that the novel *The Castaway* could be sold for on the secondary market.<sup>15</sup> In that case, *The Castaway's* publisher, the Bobbs-Merrill Company, added a notice to each copy of the book that no dealer was "authorized" to sell the book for less than \$1. When the Straus brothers purchased a number of copies and decided to sell them for less than \$1, Bobbs-Merrill sued to enforce its \$1 price floor. Ultimately, the US Supreme Court ruled that Straus did not need "authorization" from Bobbs-Merrill (or anyone else) to sell the books at whatever price they chose. Once Bobbs-Merrill sold the books, their preferences for how the books were used did not matter.

This hostility toward the secondary market continued even after the Bobbs-Merrill decision tamped down publishers' legal power to limit it. In 1931, a group of book publishers hired PR pioneer Edward Bernays – the "father of spin" – to fight against used "dollar books" and the general practice of book lending. Bernays decided to run a contest to "look for a pejorative word for the book borrower, the wretch who raised hell with book sales and deprived authors of earned royalties." The contest generated an impressive list of verbal assaults on those who would dare to lend or receive a book without paying for the privilege to do so. Suggested names included "book weevil," "bookbum," "culture vulture," "book-bummer," and "book buzzard."<sup>16</sup>

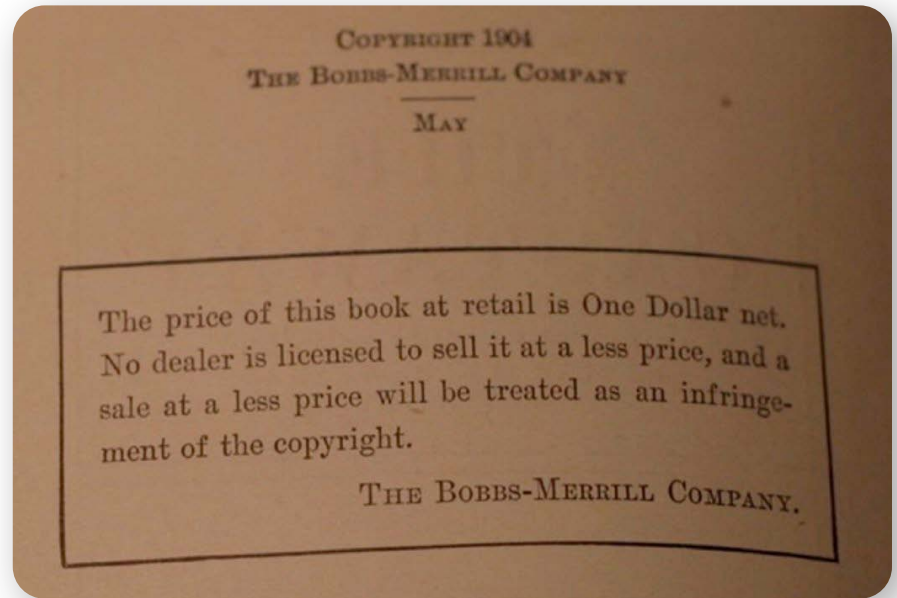
<sup>15</sup> *Bobbs-Merrill Co.*, 210 U.S. at 341.

<sup>16</sup> "Booksneak" emerged victorious, although it does not appear to have entered widespread use. Ted Striphas, *The Late Age of Print: Everyday Book Culture from Consumerism to Control* 35 (2009).



**Figure 4. Bobbs-Merrill Company v. Straus and the establishment of the First Sale Doctrine.**

The Bobbs-Merrill Company added a notice to each copy of *The Castaway* that no dealer was “authorized” to sell the book for less than \$1. Ultimately, the US Supreme Court ruled that Straus did not need “authorization” from Bobbs-Merrill (or anyone else) to sell the books at whatever price they chose.



In the introduction to C. E. Ferguson’s 1969 revised-edition economics textbook, he genially declared that “[s]ince everyone knows the basic reason for a revised edition is to kill off the existing used book market, it would be idle to suggest otherwise.”<sup>17</sup> Similarly, in 1974, it was reported that “[p]ublisher’s representatives have no doubt” second-hand copies of textbooks acted as “a drain on publishers’ and authors’ profits.”<sup>18</sup>

Publishers demonstrated their antipathy toward the secondary market again when Amazon launched its secondary book market. In 2002, Association of American Publishers (AAP) President Patricia Schroeder told *The New York Times*: “I wring my hands, pound my desk and say, ‘Aargh,’” because publishers could not legally prevent Amazon from immediately offering used books for sale (at a price outside of the publishers’ control), once those books entered the market.<sup>19</sup> Schroeder pleads for more control over the secondary market: “We asked could we at least talk about when something could become available as a used book? Could we maybe wait three months after the book was published?”

<sup>17</sup> C. E. Ferguson, *Microeconomic Theory* at vii (rev. ed. 1969).

<sup>18</sup> H. Laurence Miller Jr. On Killing off the Market for Used Textbooks and the Relationship Between Markets for New and Secondhand Goods, 82*J. Pol. Econ.* 612, 612 (1974).

<sup>19</sup> David D. Kirkpatrick, Online Sales of Used Books Draw Protest, *NY Times* (Apr. 10, 2002), <https://www.nytimes.com/2002/04/10/business/online-sales-of-used-books-draw-protest.html>

Amazon founder Jeff Bezos responded to the complaints by reminding publishers about the First Sale Doctrine: “When someone buys a book, they are also buying the right to resell that book, to loan it out, or to even give it away if they want. Everyone understands this.”<sup>20</sup> In short, Bezos said tough luck – ownership confers certain rights, including the right to resell.<sup>21</sup>

Having no luck in judicial courts or the courts of public opinion when it came to controlling the secondary market for physical books, publishers saw a new opportunity as they began to offer ebooks to the public. They saw ebook licenses – not sales – as a way to retain control over digital books by preventing secondary markets from developing. Platforms like Amazon, and other secondhand booksellers, cannot resell ebooks they or their users do not own. Borrowing from the software industry’s playbook, publishers adopted the practice of licensing content from ebooks instead of selling ebooks outright. By licensing ebooks, publishers could claim they avoid transferring digital ownership, which would create first sale rights for purchasers. A research study participant formerly at a Big Five publishing house explained the practical effect of this approach: “With a license, you can’t give it away. You can’t sell it. You can use it. And you can see why you have to do that. Right? Because if we sold you the digital file, you could resell it to whoever you want at whatever price you wanted. And there’s no limit.”

Although concerns about the secondary market usually are discussed in the context of used book sales, participants in the publishing industry also expressed wariness about book owners lending their books to friends. As a former executive at a major publishing house noted, one benefit of a license was that once you pay for an ebook, “you can’t give it away.”

Decades of publisher hostility does not prove that the secondary market actively harms the interests of publishers or authors. While there are publishing employees who do not seem to believe the resale of books is an ongoing concern to publishers, because they are resold through bookstores and similar venues, there is still evidence of a pervasive belief within the publishing industry that the ability of book purchasers to resell their books outside of publisher control is a problem that needs to be addressed. This belief may help explain the publishing industry’s enthusiasm for digital distribution models that eliminate first sale and the independent secondary market entirely.

---

20 Rob Beschizza, The Point of Kindle Is to Burn Books, *WIRED* (Nov. 21, 2007), <https://www.wired.com/2007/11/the-point-of-ki/>

21 It’s worth noting that, just five years later, when Amazon introduced its Kindle digital ebook reader, the device’s Terms of Service said the opposite about the company’s digital books. Instead of conferring full ownership, Amazon told users that they couldn’t “sell, rent, lease, distribute, broadcast, sublicense or otherwise assign any rights to the Digital Content or any portion of it to any third party.” *Kindle Store Terms of Use*, Amazon, <https://www.amazon.com/gp/help/customer/display.html?nodeId=201014950> (last visited Mar. 27, 2023).



## Fear, Loathing, and Imitation of the Music Industry’s Digital Marketplace

The vendors and publishing workers interviewed for this study consistently cited the history of the music industry’s trials, tribulations, and ultimate transition to digital as a cautionary tale that helps explain why publishers adamantly prefer to structure all ebook purchases as licenses instead of sales. Publishers and booksellers saw music industry empires crumble in the early 2000s as open digital music formats such as MP3 became popular. Companies like Sam Goody and Tower Records shuttered their physical storefronts as people started downloading music from unauthorized sources like Napster, Kazaa, Aimster, Grokster, and LimeWire instead of buying physical copies.

For most study participants, the music industry acted as a powerful negative example of what could happen during a digital transition. It was cited by the publishers as a foundation of the current ebook market in proceedings against the Internet Archive’s Controlled Digital Lending service. “Mindful of the havoc wrought on the music industry by piracy and the uncontrolled appropriation of digital content, the Publishers developed a sustainable market for affordable ebooks with critical controls, including limiting commercial ebooks to one user account.”<sup>22</sup>

Later, a senior vice president of a Hachette imprint reiterated these concerns in the same proceeding. “Many authors (and some publishers) were initially wary of entering the ebook market, in large part because it emerged in the wake of the destruction that websites like Napster caused to the music industry.”<sup>23</sup>

Chantal Restivo-Alessi, the chief digital officer of HarperCollins, serves as a personal bridge between the two industries. “The ebook was developed in the shadow of rampant piracy in the music industry in the late 1990s and early 2000s. Concerns about the widespread unauthorized distribution of digital copies were front of mind when the company first began to evaluate the ebook format decades ago and have remained in the forefront of my thoughts once I joined the company in 2012. I know this history first-hand because I worked in the record industry from 1996 to 2007.”<sup>24</sup>

While it is not entirely clear how this history has influenced individual decision-making, its pervasiveness in discussions around ebooks is noteworthy in and of itself.

---

22 Pl.’s Mem. in Supp. of Mot. Summ. J. 12, ECF No. 99 (internal citations omitted).

23 Decl. of Ben Sevier 15, ECF No. 92.

24 Decl. of Chantal Restivo-Alessi 6, ECF No. 94.

“Books are a thin-margin business,” a former executive at a Big Five publisher said. “So if you have a drop of say, 15% of people who go into bookstores, they will not be able to make it through. And so you would have what happened to music stores happen to bookstores, and nobody in publishing wants to see that.”

One study participant, who is an executive with a book trade association, believes publishers misinterpreted how the music industry was ravaged by these changes.

“There’s some aspects of what happened with music that applies to journal publishing and there are some types of books that might emulate what’s happening to journal publishing, but I think overall, what happened with music was that people didn’t want the bundle. They didn’t want the album. Most people, they wanted a song, or they wanted a couple of songs. And there was no market to create those songs until Napster came along. And Napster did it in a way that violated copyright and got it shut down but it opened the barn door to oh, we can sell songs and Apple stepped in, particularly among other [technology companies], and said, ‘Alright, we can help you, the label, sell more songs.’”

In the early 2000s, Amazon experimented with the idea of selling a few pages of a book digitally. As Jim Zarroli reported for NPR in 2005, there didn’t appear to be much of a market for this service and it didn’t take long for Amazon to discontinue this service.<sup>25</sup>

“People don’t buy chapters,” the same participant added. “They do in education, but not in trade. It’s not to say that there isn’t some Napsterization of aspects of book publishing; I think for retail it’s not really relevant.”

Nonetheless, what happened with music and other entertainment formats was enough to have publishers concerned. As one executive with an ebook platform company recalled, in the early 2010s, publishers were extremely worried about the entire industry’s ability to be turned upside down by piracy.

“In 2021, 2022, 2023, they’re looking at it more from the point of view of subscription and streaming models and the change to economics that happens when you move from selling things one at a time to selling things through subscription services. So they seem to spend a lot of time thinking about the music industry.”

The focus on unauthorized file-sharing systems like Napster as a threat to ebook publishers overshadows the eventually successful, legitimate market for digital music. Today, music labels can distribute copies of music on platforms like Apple’s iTunes and Amazon’s MP3 Store and independent music sites like Bandcamp. These sites all sell music files to consumers, yet the publishing industry appears to have ignored these models in favor of more platform-based

25 Jim Zarroli, *Amazon to Sell Full Books, or a Few Pages*, Online, NPR (Nov. 4, 2005), <https://www.npr.org/templates/story/story.php?storyId=4990112> Amazon does still maintain a “Look Inside” program that allows publishers to offer a free sample of a limited number of pages to prospective buyers. <https://www.amazon.com/gp/help/customer/display.html?nodeId=201995070>

licensing approaches such as Spotify. While the experience of individual artists has varied significantly, the fact that the music industry has managed to stay afloat despite the loss of physical storefronts and albums, benefitting from new platform streaming models that depend on user subscriptions, digital ads, and data collection, has had an impact on the publishing mentality.<sup>26</sup>

“Music went through some bumps, but now I would say the music industry is in a pretty great state, right?” an executive at an ebook subscription service platform said. “Between Spotify and Apple Music, there’s still a lot of subscription revenue out there for music. So I think there may be some bumps in the road in embracing the escription model, but I think that in the long run it’s going to be much better for the [publishing] industry for everyone to just embrace it.”

In some ways, fears that are discussed in the context of the music industry may really be anxiety about change more broadly. Multiple study participants that maintain professional relationships with major publishers recall publishers’ long-term reluctance to embrace the ebook format simply because it was outside their comfort zone.

## Publishers Interpret Changes in Copyright Law As An Opportunity to Shut Down Secondary Markets

Courts have not settled on how to treat digital ownership, and publishers are taking advantage of this unresolved legal issue. The advent of digital technology has complicated and obfuscated the distinction between sales and licenses for digital goods.<sup>27</sup> One reason that copyright is so complicated online is the way digital goods are copied. When physical books are copied, it is easy to quantify and track those copies. Each copy can be counted, lifted, stacked, and stored on a tangible shelf. Digital copies are much harder to quantify and track.<sup>28</sup> With digital technology, there is rarely, if ever, a single copy of a work, even when the work is sold. Copies bought on CD or DVD are then loaded into the memory of a laptop computer. Downloads purchased via the cloud end up stored locally on a phone or tablet and across numerous cloud servers.

26 Marisa Dellatto, Spotify Says it Paid \$7 Billion in Royalties in 2021 Amid Claims of Low Pay from Artists, *Forbes* (Mar. 24, 2022), <https://www.forbes.com/sites/marisadellatto/2022/03/24/spotify-says-it-paid-7-billion-in-royalties-in-2021-amid-claims-of-low-pay-from-artists/>

27 See, e.g., Aaron Perzanowski, Fixing RAM Copies, 104 *NW. L. Rev.* 1067 (2010); Aaron Perzanowski and Jason Schultz, Digital Exhaustion, 58 *UCLA L. Rev.* 889 (2011); Aaron Perzanowski and Chris Jay Hoofnagle, What We Buy When We Buy Now, 165 *U. PA. L. REV.* 315 (2017).

28 Plaintiffs’ expert in *Hachette v. Internet Archive* attempted to quantify all of the individual copies involved in the Internet Archive’s process of digitizing a physical book. They identified at least 11 distinct copies that existed as incremental steps in creating the final ebook, although those did not include an undetermined number of additional files created during the creation of cover art, or to make additional file versions. Supplemental Expert Report of Prof. Ian Foster at 44–46, *Hachette v. Internet Archive*, No. 1:20-CV-04160 (S.D.N.Y. July 7, 2022), ECF No.100-2.

**This rhetorical shift from “selling” individual copies on digital media to “licensing” a collection of multiple copies for use in digital memory began to obfuscate who controlled the digital item that was sold to the consumer.**



Digital copies, even those purchased lawfully, often required a temporary instance of the software to be loaded into digital memory. As a result, copyright owners began to claim that a copyright license for these additional copies was needed to cover the copies of the work that existed in computer memory.<sup>29</sup> This rhetorical shift from “selling” individual copies on digital media to “licensing” a collection of multiple copies for use in digital memory began to obfuscate who controlled the digital item that was sold to the consumer.

Even the notion of a “copy” soon became more metaphysical than physical in the world of digital content. The existence of a single “copy” became illusory, especially as modern computer storage systems learned to fragment even single files into thousands or millions of “shards” so they could be stored efficiently across various memory locations. Those same systems can make multitudes of ephemeral copies of files in the process of achieving mundane tasks like displaying the page of an ebook on a screen. Even if a consumer could argue that they owned the “copy” they bought, the tough question soon became: Where exactly is this copy? And how does the consumer find it if they want to resell it, lend it, or even give it away to a friend or family member?

## Courts Begin to Weigh In

US courts struggled with how to think about the notion of a digital copy, and how to distinguish between a digital sale and a digital license. While some early courts were able to draw a distinction by analyzing a CD or DVD sale as controlled by copyright law and the installation of the digital content on a computer governed by whether a separate “contract” existed between the seller and the buyer, later courts could not make such clear-cut distinctions between copyright, which is governed by the First Sale Doctrine, and contract law-bound ownership arrangements, which are not.<sup>30</sup>

In the summer of 2010, the US Ninth Circuit Court of Appeals heard three different copyright appeals, all of which turned on the notion of whether or not a consumer who had purchased or been gifted a digital copy could make independent uses of that copy against the wishes of the copyright owner.

<sup>29</sup> See, e.g., *Step-Saver Data Sys., Inc. v. Wyse Tech.*, 912 F.2d 643 (3d Cir. 1990).

<sup>30</sup> See *SoftMan Prods. Co. v. Adobe Sys.*, 171 F.Supp.2d 1075, 1087–89 (C.D. Cal. 2001) (differentiating between the sale of a CD with a copy of an Adobe program, which would make the purchaser the owner of that copy, and the loading of Adobe software onto computers, which would make the user subject to the terms of an Adobe license). But see *Krause v. Titleserv, Inc.*, 402 F.3d 119, 124–25 (2d Cir. 2005) (finding ownership of copies of source code even in the absence of a formal title where owner “paid substantial sums, to possess and use a copy indefinitely without material restriction, as well as to discard or destroy it at will”).

The first case was *UMG Recordings, Inc. v. Augusto*.<sup>31</sup> The major record label UMG had a practice of sending unsolicited promotional (“promo”) CDs to DJs and other music fans. Troy Augusto, an entrepreneurial CD collector, would lawfully purchase used promo CDs and resell them, even though they often had a label on them stating “Not for Resale.” The second case was *Vernor v. Autodesk*, a similar case in which Tim Vernor bought and resold used copies of Autodesk’s architectural software program, AutoCad.<sup>32</sup> Finally, in *MDY v. Blizzard*, the court heard a case as to whether purchasers of Blizzard’s *World of Warcraft* (WoW) video game infringed its copyright when they used a third-party “bot” program to autopilot their videogame characters through menial tasks in the game, a practice which Blizzard had forbidden in its game rules.<sup>33</sup>

In each of these cases, questions of digital ownership were central. Despite the similarities among the fact patterns, the same three judges decided all three cases differently. Augusto emerged victorious, the court deciding that the lawful recipient owners of the promo CDs had a legal right to give their CDs to anyone and, thus, he had a legal right to resell them as he pleased. Vernor, on the other hand, was out of luck. The court decided that the Autodesk license was somehow stronger than the “Not for Resale” label on the CDs, and that digital software, oddly, was in some way inherently different from digital music, without saying exactly how. And in MDY, the court went a third way, saying that depending on the license term and language, the consumer sometimes had a right to use the software as she liked, and other times, she didn’t. It appeared that the terms copyright holders attached to digital media were important, although it remained unclear exactly how the power of terms interacted with the rights of users.

Then, in 2018, the Second Circuit Court of Appeals gave its own gloss on digital ownership. In *Capitol Records v. ReDigi*, the court considered whether a consumer who “bought” digital music songs from Apple’s iTunes Store had a right to resell them, and relatedly, whether the online marketplace ReDigi could help facilitate such sales.<sup>34</sup> Here, the battle over rhetoric, technology, and copyright law reached a fever pitch. On the one hand, ReDigi had engineered a system that ensured the transfer of a music file securely from the seller’s computer to the purchaser’s without any extra copies being made. On the other hand, because the transfer was done through ReDigi’s cloud-based service, the transfer process involved making multiple intermediary copies of the original file, including the final copy on the purchaser’s computer. In deciding the case, the court interpreted this second digital location as a “copy” of the file, instead of a relocation of the original file from one computer to the other. And as such, that made the buyer’s copy an infringing reproduction of the song instead of a resale of the buyer’s copy.

---

31 *UMG Recordings, Inc. v. Augusto*, 628 F.3d 1175 (9th Cir. 2011).

32 *Vernor v. Autodesk, Inc.*, 621 F.3d 1102 (9th Cir. 2010).

33 *MDY Indus., LLC v. Blizzard Ent., Inc.*, 629 F.3d 928 (9th Cir. 2010).

34 *Capitol Records, LLC v. DeRigi Inc.*, 910 F.3d 649 (2d Cir. 2018).

All of this conflicting case law left quite a state of confusion for what it means to “own” digital goods. To some degree, that is still the case. The Supreme Court, to its credit, has tried to clarify in both a copyright case involving internationally sold used books and a patent case involving the reselling of printer toner cartridges that the First Sale Doctrine means just that – copyright and patent owners get to set the terms only the first time they sell a copy that contains their work.<sup>35</sup> But, because neither case was entirely digital in nature, the conflict and confusion between the Supreme Court’s approach and cases like *ReDigi*, *MDY*, *Vernor*, and *Augusto* remains.<sup>36</sup>

The vast majority of publishers have since taken advantage of this confusion, and have begun to insist that approaches like *Vernor*, which maximize the control that license terms can have on users, represent the current state of the law.<sup>37</sup> What the US courts or Congress will ultimately decide on these issues and how that will impact digital ownership of ebooks remains to be seen. In the meantime, publishers continue to move forward on building a walled garden for ebooks built on licenses and technological limitations for consumers and cultural institutions.

## Digital Distribution Changes the Economics of Book Production

Another murky area of the ebooks analysis is what it costs to publish an ebook. Is it more expensive to publish paper or electronic copies of books? Publishers sometimes claim that it is expensive to publish an ebook, but concrete numbers about how much ebooks cost to produce are hard, if not impossible, to find. Although reliable empirical data is limited, ebooks appear to be relatively inexpensive to produce when compared to physical books. One participant from a Big Five publisher stated simply that “Ebooks are easy to create in a form that everyone accepts as acceptable.” What that means, according to one digital sales representative at a Big Five publisher, is that the people who work on ebook files have to “design the file in such a way that it recognizes what device it’s being rendered on. We make sure that [the file] can be rendered on that device.”

This individual went on to explain that when file formats are not fully compatible across devices, text can display with unwanted spaces and characters that take away from the user’s experience. “When a book is designed, believe it or not, it’s like a work of art. We actually have book designers who actually design how fonts look on a page.”

---

35 *Kirtsaeng*, 568 U.S. at 530; *Impression Prods. v. Lexmark Int’l, Inc.*, 137 S. Ct. 1523 (2017).

36 A recent ruling in favor of publishers in *Hachette v. Internet Archive* may further complicate this analysis. The district court found that the Internet Archive’s Controlled Digital Lending program was not protected by fair use. The Internet Archive used that program to lend out scanned digital copies of physical books in its collection.

37 Michelle C. Sisto, *Publishing and Library E-Lending: An Analysis of the Decade Before COVID-19*, 38 *Publ’g Rsch. Q.* 405.



**“A lot of our effort in print is dedicated toward optimizing that in terms of, you know, the paper that we acquire, the printing, space, transportation, shipping. None of that is required for ebooks.”**



A participant from a trade publishing association had a more nuanced view that comes across as more critical of the process: “Generally people see digital books as more expensive, but I would argue that they are not in terms of prep work, and that when it’s more expensive it’s because your workflow wasn’t built to support multiple outputs.... I’m keenly aware of what happens with digital formats when you first start trying to do them. You create a separate workflow with separate processes, you make more mistakes, things don’t work out, and then they kind of become calcified.”

According to an independent publisher, “If it’s a book with text, it just means creating a flowable epub file, which is not hard. If it is an illustrated book, then a fixed format epub is usually created, unless the publisher wants to pay extra for fancy features and custom coding.” Flowable epub files work across platforms because they are responsive and the text works and “reflows” in whatever device it’s being used on.<sup>38</sup>

The relative economy of ebooks may also flow from the fact that ebook files are created once at the time of initial publication and then distributed in response to demand. One Big Five publishing house employee noted the inefficiencies inherent in the physical book distribution model when compared to ebooks: “A lot of our effort in print is dedicated toward optimizing that in terms of, you know, the paper that we acquire, the printing, space, transportation, shipping. None of that is required for ebooks.”

The economy of ebooks can also make hits especially profitable when, as a study participant from a trade association confirmed regarding ebooks, “the marginal cost is closer to zero.” A former Big Five executive also seemed to confirm this thought: “In an ebook, there’s no variable cost at all once you set up the system, right? You can put one book through, 100 books through, 50,000 books through. No additional cost, you’re just transferring files.”

Even if a book is a bust, there are still cost savings when publishers are dealing with the ebook format. In ebook publishing, there are no unsold copies. Publishing industry norms allow bookstores to return any unsold copies of a book to the publisher. In those cases, the publisher has paid to create the book, ship the book to the bookseller, return the book to the publishers, and destroy the copy. These costs do not exist when ebooks are distributed on demand.




---

38 <https://www.thebookdesigner.com/going-against-the-flow-reflowable-vs-fixed-layout-ebooks/>

# Ebooks Reduce the Friction in Library Lending, Raising Concerns for Publishers



In addition to music industry-related fears, publishers are afraid that libraries lending ebooks will undermine the publishing business in a way that paper book lending doesn't. Although some argue that ebook licensing helped increase access to broad digital collections during the COVID pandemic, ebooks and ebook licensing also limit the content and distribution of libraries' collections.

In addition to visiting libraries to pick out paper books to take home, library patrons also use their library cards to access ebooks through websites that are made by third-party vendors and accessible on library websites. Instead of owning or controlling books, libraries sign digital library "access agreements" with companies like Hoopla, which select which ebooks library patrons can access and decide the terms of that access and lending for the libraries.<sup>39</sup> These types of arrangements with ebook platforms worry library workers, who feel their roles are being replaced by the new vendors, or eliminated entirely. Other ebook platforms like OverDrive do let librarians handle decisions around collection development for ebooks because libraries can pick and choose individual titles. However, the same cost and restrictions still apply to librarians that use OverDrive.

---

<sup>39</sup> See, e.g., *OverDrive Digital Reserve Access Agreement*, OverDrive (Jan. 2023), <https://company.overdrive.com/dlr-aa.pdf>



## Publishers Worry That Patron Lending May Undermine Sales

Publishers worry that one-click access through libraries will discourage readers from purchasing books. One participant from a Big Five publisher in digital explained, “What is the difference between, you know, say, there’s a new book coming out, and a customer has a choice to buy it on Amazon or log on at home and download it from their library for free? There’s much more incentive to download it for free and if there’s no limit to whether that book is available.... we don’t want everyone who wants to read [the] book knowing that they can get it through the library the same day and date of the release and they don’t sell anything at all.”

Publishers made a similar point in their lawsuit against the Internet Archive’s Controlled Digital Lending program. “The convenience of library ebook lending – particularly the ability to check out ebooks instantaneously, anytime and anywhere – makes it a highly attractive proposition to consumers who might otherwise purchase ebooks.”<sup>40</sup>

This concern is often described as a lack of “friction” compared to physical book lending. It is rooted in the view that the ease of accessing books on ebook platforms like OverDrive’s Libby is now roughly equal to the ease of purchasing those same ebooks.

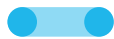
Participants associated with the publishing industry often drew a contrast between the similarity of the ebook purchasing and lending experiences on one hand, and the differences between the physical book purchasing and lending experiences on the other. However, those comparisons often exclusively attributed burdens to lending physical books that exist for purchasing physical books as well. For example, travel time and the possibility of a book not being available was often cited as a point of friction in library lending, while being essentially ignored for bookstore purchases.

A former executive at a Big Five publishing house used this framework to describe the friction involved with lending physical books: “If I’m a consumer in the old days, I could buy a book in the bookstore and read it. Or I could get in the car, drive down to the library instead of the bookstore, look to see if they had it. Check it out. If I hadn’t finished it in two weeks, I was going to have to bring it back. I was gonna forget to bring it back. I was gonna have to pay a fine. Oh, my kid lost it. Where the hell did that book go? Now I’m gonna get charged for the goddamn book and I don’t even have it.” While the post-check-out friction is specific to libraries, everything up to that point would also apply to the bookstore experience.

---

40 Pl.’s Mem. in Supp. of Mot. Summ. J. 13–14, ECF No. 99.

**“For the consumer, borrowing a free ebook from a library’s website, on the one hand, or obtaining it from Amazon or other e-retailers, on the other hand, are extremely similar experiences, except that one is free and one is not.”**



This concern was echoed by the senior VP of Grand Central Publishing (an imprint of Hachette) in the Internet Archive Controlled Digital Lending litigation. “We believe that the ‘friction’ involved in checking out physical books from libraries – the delays when a popular book is not available, the time and costs associated with traveling to the library to check out a physical book – may motivate some library patrons to purchase their own copy rather than check the book out from the library.”<sup>41</sup>

While participants from the publishing industry often drew sharp contrasts between the lending and purchasing experience of physical books, they highlighted the similarity between lending and purchasing ebooks. The same executive continued: “Perhaps most importantly, the convenience of ebook lending – particularly the ability of users to check ebooks out instantaneously anytime and anywhere – makes it a highly attractive proposition to consumers who would otherwise purchase books instead of going to the library. Critically, the experience of checking out a library ebook is functionally identical to downloading a copy from Amazon.”<sup>42</sup>

The chief digital officer at HarperCollins drew the same connection. “For the consumer, borrowing a free ebook from a library’s website, on the one hand, or obtaining it from Amazon or other e-retailers, on the other hand, are extremely similar experiences, except that one is free and one is not.”<sup>43</sup>

A senior VP at Wiley expanded: “An ebook would be capable of far more efficient reproduction and distribution than a print book, allowing it to serve exponentially greater demand.”<sup>44</sup>

Publishers are also concerned that library lending of ebooks will “train” readers that the value of ebooks is zero. This fear of training in the ebook market appears to be more acute than in the physical book market because of the publisher perception that ebook lending and purchasing is more similar than physical book lending and purchasing. The same former publishing executive explained: “And then the final significant danger is the deterioration in value. And that comes primarily through library lending. Right? The library lending issue is very difficult because it educates a whole group of people that books are essentially free.”

41 Decl. of Ben Sevier 19, ECF No. 92 (internal numbering omitted).

42 Decl. of Ben Sevier 20, ECF No. 92 (internal numbering omitted).

43 Decl. of Chantal Restivo-Alessi 11–12, ECF No. 94.

44 Decl. of Alan Pavese 9, ECF No. 93.

This concern about “free” books is echoed in the publishers’ complaint against the Internet Archive’s Controlled Digital Lending system: “[B]y providing copies of digital books for free, [the Internet Archive] devalues the book market. Consumers begin to view works as cheap and become increasingly unappreciative of what it takes to produce them and unwilling to pay fair value for them.”<sup>45</sup> This concern is not unprecedented – music platforms have been similarly critiqued for devaluing the artists’ and labels’ labor.<sup>46</sup>

The former publishing executive worries that nobody will buy ebooks if they can get an identical copy for free from a library. “And that lowers the amount of cash that comes into the system, which lowers the author’s advances, which makes it more difficult to be a publisher, and eventually makes it very difficult to make a living writing books.”

In conversations with participants, it was not always clear why publishers tend to view library copies of ebooks as “free.” Although these copies are free to patrons, they are available to patrons because libraries have purchased them from publishers. Increased patron demand for individual titles would encourage libraries to purchase additional copies from publishers. Publishers highlight that library acquisitions contribute to profits in their complaint against the Internet Archive’s Controlled Digital Lending program: “Harm to and loss of community support for public libraries, in turn, hurts the publishers, since these libraries pay for their books, which, in turn, hurts authors, who share in and depend upon compensation for their copyrights.”<sup>47</sup>

In that same proceeding, one executive touted the growth of revenue in the library ebook market. “In fiscal year 2015, library ebook licenses in the United States constituted 2.3% of HarperCollins’ total US ebook revenue. By fiscal year 2020, that share grew substantially to 7.4% and reached 12.9% of all US ebook revenue by mid-2021. (HarperCollins’ fiscal year ends June 30.) Indeed, as of mid-2021, HarperCollins earned over \$19 million in revenue from library ebook licenses in the United States, which was nearly twice as much revenue as we earned from library ebook licenses in the United States during FY 2020.”<sup>48</sup>

It is not entirely clear how publishers understand the relationship between library acquisitions and retail sales. In a declaration that is part of the Internet Archive Controlled Digital Lending litigation, one publishing executive described conclusions he drew from OverDrive circulation data (the data and analysis itself was filed under seal). “While library ebook circulations made up 50% of the total units downloaded of trade ebooks between 2017 and 2020, ebooks obtained by libraries which resulted in those circulations brought in only 13% of total ebook revenue

---

45 Although the publishers’ suit is centered on the way the Internet Archive acquired the ebooks used in the Controlled Digital Lending system, in this context the publishers are expressing concern that the patrons are not paying for the ebooks themselves. Compl. 45, ECF No. 1.

46 Hugh Zehr, *An Economic Analysis of the Effects of Streaming on the Music Industry in Response to Criticism from Taylor Swift*, 23 *Major Themes in Economics* 51, 51–63 (2021).

47 Compl. 46, ECF No. 1.

48 Decl. of Chantal Restivo-Alessi 16–17, ECF No. 94.

for Hachette.... The data discussed above has led Hachette to conclude that readers who were previously purchasing retail ebooks are switching to library ebooks, which means that we lose retail sales to free borrows.<sup>49</sup> Left unexplored in the publishers' public analysis was how many of those library loans they believed replaced potential retail sales, how many loans per ebook publishers would view as reasonable, and how many loans per ebook libraries expect when deciding to pay for ebook access in the first place.<sup>50</sup>

In one podcast, a Macmillan executive alleged that people used libraries to get around buying books: "I have a friend right now who has 11 library cards, never waits for a bestselling book, used to spend \$500 a year on books."<sup>51</sup> Study participants affiliated with the publishing industry were often quick to point to that type of lost revenue when discussing library ebook lending in interviews. There was much less discussion of the impact that this type of patron ebook consumption might have on the acquisition decisions and budgets of the libraries used by those patrons.

## Ambiguity Around Degradation Rates of Paper and Ebooks

Another concern that publishers shared was the fear that each electronic book copy is indestructible compared to paper copies of the past. Participants from the publishing industry often noted during our interviews that paper books tend to degrade physically, especially when they are shared.<sup>52</sup> This was often contrasted with ebooks, which were commonly described as lasting in perpetuity.<sup>53</sup> This industry understanding of the relative durability of books and ebooks – regardless of its veracity – raises questions about how to think about the value of a single copy of a book, in either physical or electronic form.

---

49 Decl. of Ben Sevier 21–22, ECF No. 92 (internal citations omitted).

50 Current library licensing practices, which often limit the total number of lends over a period of time, do provide some additional insight into publisher views. However, the calculations underpinning the limits embodied in the licenses remain opaque.

51 Planet Money, *The E-Book Wars*, NPR (Nov. 11, 2022), <https://www.npr.org/transcripts/1135639385>

52 Although HarperCollins linked its 2011 26-lend limit to an estimate of the average life expectancy of a book on a shelf, that number was immediately questioned. Carolyn Kellogg, HarperCollins' 26-checkout limit on libraries' ebooks starts today, *Los Angeles Times* (Mar. 7, 2011), <https://www.latimes.com/archives/blogs/jacket-copy/story/2011-03-07/harpercollins-26-checkout-limit-on-libraries-ebooks-starts-today>

53 For example, the publisher's complaint in the recent Controlled Digital Lending lawsuit against the Internet Archive asserts that ebooks "do not degrade over time as physical books do." Compl. 6, ECF No. 1. This claim was repeated later in the proceedings: "And, because ebooks do not get lost or damaged, a library would not have to organically repurchase a title once it has paid for an electronic copy." Decl. of Ben Sevier 20, ECF No. 92 (internal numbering omitted).

The physical fragility of paper books created an opportunity for publishers to resell copies of a popular title multiple times, especially to libraries, where books go through a lot of wear and tear as they are read by numerous patrons. As a former Big Five publishing executive explained, “Physical books fall apart. And they actually fall apart reasonably fast. Physical books get lost. Physical books don’t get returned to the library. So the library has to keep ordering the book every once in a while; think how many times your local library has ordered up *Pride and Prejudice*. Right? Since it was published. They’ve ordered a lot of copies of that book.”

The same former executive contrasted the fragility of physical books with ebooks: “Digital file, oh, that lasts forever. They buy it once, it lasts forever.” Another executive simply asserted that an ebook “does not degrade.”<sup>54</sup>

While this sentiment was common among participants, it is unclear how accurately it reflects real-world dynamics. Popular physical books may wear out quickly, but copies of other titles remain in a library’s collection for years, decades, or even centuries, waiting for the occasional patron interest. When it comes to digital, publishers have demonstrated that they can create licensing schemes that require libraries to re-purchase books at a regular interval, or after a certain number of patrons check them out.<sup>55</sup>

Furthermore, as anyone who has ever tried to open a digital file that is more than a few years old may know, the “forever” nature of digital copies is far from proven. The Internet Archive has described the work involved in maintaining its digital collection, in that it “processes and reprocesses the books it has digitized as new optical character recognition technologies come around, as new text understanding technologies open new analysis, as formats change from djvu to daisy to epub1 to epub2 to epub3 to pdf-a and on and on. This takes thousands of computer-months and programmer-years to do this work.”<sup>56</sup>

It may be relevant that the nature of this degradation is not identical between physical books and ebooks. Physical books tend to wear out in a way that correlates with their popularity. The most popular titles wear out quickly, while less-popular titles may last in perpetuity (or at least until the acidic compounds in the paper they are printed on destroys them). In contrast, ebooks tend to become less accessible over time. All titles grow obsolete at the same rate, and all may disappear if the platform distributing them goes out of business. As that happens, the most popular titles may be the only ones that are replaced.

Regardless of their rate of decay, some libraries may devote resources to maintaining their digital collections, just as some libraries may devote resources to maintaining their physical collections. However, other libraries could adopt the same strategy for replacing outdated ebooks that they already use for worn-out physical books: purchase new copies.

---

54 Decl. of Chantal Restivo-Alessi 11–12, ECF No. 94.

55 Maribel Riaza and Javier Celaya, *A Guide to eBook Licenses Purchase Models for Libraries and Publishers* (2015), <https://www.dosdoce.com/wp-content/uploads/2015/11/Guide-to-Ebook-Licensing-Purchase-Models-Bookwire-and-Dosdoce.pdf>

56 Brewster Kahle, *Digital Books Wear out Faster than Physical Books*, Internet Archive (Nov. 15, 2022), <http://blog.archive.org/2022/11/15/digital-books-wear-out-faster-than-physical-books/>

## Concentration and Consolidation Make It Easier to Push the Market in Publishers' Favor

The changes in copyright law and technological realities have been unfolding against a backdrop of continued concentration within the publishing industry itself. This consolidation of power among publishers made it easier to avoid competitive pressures that may have prevented the ebook market from evolving in a way that is so different from the physical book market. It has also given publishers leverage in the ebook market to force libraries to follow their lead. This leverage is most vividly illustrated by examining the history of the terms that the largest publishers have offered to libraries.<sup>57</sup>

Initially, ebook terms for libraries were relatively close to the rules that govern the first sale of physical books. As early as 2001, HarperCollins was offering libraries unlimited-checkout licenses to ebooks.<sup>58</sup> While not quite a sale, this unlimited model looks remarkably flexible in retrospect.

2011 marks a shift in terms for many of the larger publishers. Citing “security concerns,”<sup>59</sup> Penguin stopped offering new ebook titles via the OverDrive platform used by many libraries. Hachette also stopped offering frontlist ebooks to libraries, while also doubling prices for certain backlist titles.<sup>60</sup> HarperCollins began capping ebook licenses at 26 loans.<sup>61,62</sup> Once an

57 These publishers were once known as the Big Six (Hachette, HarperCollins, Macmillan, Penguin, Random House, and Simon & Schuster), more recently the Big Five (with the merger of Penguin and Random House in 2013) and possibly someday the Big Four (in late 2022, a US District Court judge blocked an attempted merger between Penguin Random House and Simon & Schuster).

58 Randall Stross, Publishers vs. Libraries: An E-Book Tug of War, *NY Times* (Dec. 25, 2011), <https://www.nytimes.com/2011/12/25/business/for-libraries-and-publishers-an-e-book-tug-of-war.html>

59 Andrew Albanese, Citing “Security Concerns” Penguin Pulls New Titles from OverDrive, *Publishers Weekly* (Nov. 21, 2011), <https://www.publishersweekly.com/pw/by-topic/industry-news/trade-shows-events/article/62877-ala-2014-raising-the-stakes.html>

60 Andrew Richard Albanese, *ALA 2014: After Success on E-Book Access*, ALA Looks Ahead, *Publishers Weekly* (June 13, 2014), <https://www.publishersweekly.com/pw/by-topic/industry-news/trade-shows-events/article/62877-ala-2014-raising-the-stakes.html>

61 AM. LIBR. ASS’N, *The 2012 State of America’s Libraries 20* (2012) [hereinafter 2012 ALA Report], <https://www.ala.org/news/sites/ala.org.news/files/content/StateofAmericasLibrariesReport-2012Finalwithcover3.pdf>

62 The chief digital officer of HarperCollins (who joined the company in 2012) explained the motivation behind this shift as a response to the growing popularity of ebooks: “In the early 2000s, when HarperCollins first introduced its initial library ebook lending model, ebooks were read predominantly on desktop computers and the sales across all channels were a fraction of one percent of the company’s revenue. But by 2011, ebook sales had exploded, resulting in a rapid acceleration of ebook retail and library sales. A ‘perpetual’ term was a very long time for a product like an ebook that does not degrade, and our pricing of ebooks was not consistent with the number of readers who would read, respectively, print and ebook formats.” Decl. of Chantal Restivo-Alessi 11–12, ECF No. 94.



ebook had been lent 26 times, a library could re-purchase the book at a discount.<sup>63</sup> Macmillan and Simon & Schuster continued not to offer ebooks to libraries under any terms.

In 2012, Random House introduced a 100%-200% library ebook price increase on unlimited-use licenses.<sup>64</sup> While this price increase did not come with an embargo that held new titles from library circulation, it did come with a request that libraries share data about patron borrowing patterns.<sup>65</sup>

These license terms were adopted by Hachette in 2013,<sup>66</sup> with library prices three times the consumer retail price for the first year after a book's release.<sup>67</sup> That same year, Macmillan began a pilot program to license its books to libraries, offering a license that was good for the lesser of two years or 52 loans.<sup>68</sup> Simon & Schuster began its own library pilot program, which included a one-year license and a requirement that libraries act as sales agents by offering a "buy it now" option alongside ebook lending options (this "buy it now" requirement was removed by the end of 2014).<sup>69</sup>

By 2020, all of the Big Five had shifted to some form of time-limited licenses (HarperCollins in 2011,<sup>70</sup> Simon & Schuster<sup>71</sup> and Macmillan<sup>72</sup> in 2013 – followed by a shift back toward a limited perpetual approach in 2019<sup>73</sup> and a reintroduction of time limits in 2020<sup>74</sup>, Penguin

---

63 *Id.*

64 *Id.* at 17.

65 *Id.* at 18.

66 Andrew Albanese, Hachette Makes Full E-Book Catalog Available to Libraries, *Publishers Weekly* (May 1, 2013), <https://www.publishersweekly.com/pw/by-topic/digital/content-and-e-books/article/57049-hachette-makes-full-e-book-catalogue-available-to-libraries.html>

67 <https://americanlibrariesmagazine.org/wp-content/uploads/2016/01/Big-Five-Ebook-Terms-011816.pdf>

68 Jim Milliot, Macmillan to Begin E-Book Library Lending Pilot, *Publishers Weekly* (Jan. 24, 2013), <https://www.publishersweekly.com/pw/by-topic/digital/retailing/article/55630-macmillan-to-begin-e-book-library-lending-pilot.html>

69 S&S Offers E-Book Lending, Purchase Via New York City Libraries, *Publishers Weekly* (Apr. 15, 2013), <https://www.publishersweekly.com/pw/by-topic/industry-news/libraries/article/56826-s-s-offers-e-book-lending-purchase-via-new-york-city-libraries.html>; Robert C. Maier, Big Five Publishers and Library Lending, *AM. LIBRS.*, <https://americanlibrariesmagazine.org/wp-content/uploads/2016/01/Big-Five-Ebook-Terms-011816.pdf> (Jan. 18, 2016).

70 2012 *ALA Report*, *supra* note 39.

71 *S&S Offers E-Book Lending, Purchase Via New York City Libraries*, *supra* note 46.

72 Milliot, *supra* note 45.

73 John Sargent, Letter to Macmillan Authors, Macmillan Illustrators, and Agents (July 25, 2019) [https://www.publishersweekly.com/binary-data/ARTICLE\\_ATTACHMENT/file/000/004/4222-1.pdf](https://www.publishersweekly.com/binary-data/ARTICLE_ATTACHMENT/file/000/004/4222-1.pdf)

74 *Important Updates from Penguin Random House and Macmillan US*, OverDrive (Mar. 18, 2020), <https://company.overdrive.com/2020/03/18/41849/>

Random House in 2018,<sup>75</sup> and Hachette in 2019<sup>76</sup>). These terms were slightly modified during the early phase of the COVID pandemic,<sup>77</sup> although it is unclear how and if they will persist going forward.

This evolution away from licenses that simulate ownership toward licenses that simulate ongoing rentals is especially noteworthy in light of the constant pushback publishers received from one of the license's primary markets: libraries and librarians. The American Library Association “expressed concern” when Hachette moved away from perpetual licenses in 2019.<sup>78</sup> ReadersFirst described Macmillan’s reintroduction of library embargoes for new titles as “A Giant Leap ... Backwards” (ellipsis in original).<sup>79</sup> Librarian Sari Feldman’s Publishers Weekly column on the same topic described it as “A Dystopian Twist.” A viral blog post by Sarah Houghton, director for the San Rafael (California) Public Library, compared the “dismal ebook situation in libraries” to a bad boyfriend.<sup>80</sup>

None of this activity has pushed any of the Big Five to compete by adopting the model that a significant number of their customers appear to demand. Instead, over the course of almost two decades, these publishers have moved toward models that give them more control over how ebooks are used, and more opportunities to shift toward recurring revenue streams.



- 
- 75 Andrew Albanese, Penguin Random House Changes Library E-Book Lending Terms, *Publishers Weekly* (Sep. 4, 2018), <https://www.publishersweekly.com/pw/by-topic/industry-news/libraries/article/77904-penguin-random-house-changes-its-library-e-book-terms.html>
- 76 Press Release, Amer. Libr. Ass’n, *ALA ‘Concerned’ over Hachette Book Group Ebook and Audio Book Lending Model Changes* (June 17, 2019), <https://www.ala.org/news/press-releases/2019/06/ala-concerned-over-hachette-book-group-ebook-and-audio-book-lending-model>
- 77 See, e.g. Andrew Albanese, Macmillan Abandons Library E-book Embargo, *Publishers Weekly* (Mar. 17, 2020), <https://www.publishersweekly.com/pw/by-topic/industry-news/libraries/article/82715-macmillan-abandons-library-e-book-embargo.html>
- 78 Press Release, Amer. Libr. Ass’n, *ALA ‘Concerned’ over Hachette Book Group Ebook and Audio Book Lending Model Changes* (June 17, 2019), <https://www.ala.org/news/press-releases/2019/06/ala-concerned-over-hachette-book-group-ebook-and-audio-book-lending-model>
- 79 Michael Blackwell, *A Giant Leap ... Backwards*, *ReadersFirst* (July 6, 2018), <https://www.readers-first.org/news/2018/7/6/a-giant-leap-backwards>
- 80 Andrew Richard Albanese, ALA 2014: After Success on E-Book Access, ALA Looks Ahead, *Publishers Weekly* (June 13, 2014), <https://www.publishersweekly.com/pw/by-topic/industry-news/trade-shows-events/article/62877-ala-2014-raising-the-stakes.html>; Sarah Houghton, *I’m Breaking up with Ebooks (and You Can Too)*, *Librarian in Black* (Aug. 1, 2012), <https://web.archive.org/web/20120807000044/http://librarianinblack.net/librarianinblack/2012/08/ebookssuckitude.html>



# The Publisher-Platform Partnership Shapes the Ebook Market



The physical book market is largely shaped by balancing the interests of two stakeholders: publishers and purchasers. While bookstores play an important part in that market, the First Sale Doctrine tends to limit their permanent influence: A bookstore's relationship with a physical book ends the moment that it is sold. Deciding to shop at a new bookstore does not involve re-purchasing all of the books you purchased from others.

Balancing interests in the ebook market is more complex. In addition to publishers and readers, ebook platforms have ongoing interests in how ebooks are created, distributed, and controlled. Those interests exist independently from the interests of publishers and readers. Licensing allows ebook platforms to exercise that influence in an ongoing manner.

In many (although not all) cases, publishers and platforms find their interests aligned when it comes to shaping the ebook market. This publisher-platform partnership rebalances the playing field in the ebook market, tilting things away from purchasers and toward the publishers and platforms.

The result is an ebook market that is designed to support the preferences of publishers and platforms, at the expense of the interests of the purchasers. It replaces sales and ownership with licenses tying ebooks to both publishers and platforms in perpetuity. These perpetual ties also create new opportunities for publishers and platforms to monetize reader data by spying on ebook readers.

While locking purchasers into platforms serves some interests of both publishers and platforms, that lock-in has unique impacts on publishers, platforms, and purchasers.

## Digital Rights Management Technology Creates Lock-In

Licenses are the legal tools that platforms and publishers use to control how purchasers can use ebooks. Digital Rights Management (DRM) is the primary technical tool used to implement that control, sidestepping first sale rights. Publishers and book platforms use DRM to prevent consumers from downloading, saving, sharing, and other common first sale-permitted activities. DRM tools make it very hard, if not impossible (and possibly illegal),<sup>81</sup> for users to move the ebooks in their collection between platforms and ereaders.<sup>82</sup> If a consumer licenses access to a book on Kindle, they cannot open that file on a NOOK, or on any other platform or app.

The full motivations behind implementing DRM are not always entirely clear. Many participants connected DRM with concerns about large-scale piracy.<sup>83</sup> At the same time, they also noted approvingly that DRM would prevent casual lending between friends. (One publishing executive even stated that DRM played an important role in ensuring that local library spending benefited local members of the community.)<sup>84</sup>

While sometimes conflated by participants, stealing books and sharing books are distinct activities in the world of physical books and first sale. Printing and distributing multiple, unauthorized copies of a physical book would violate the copyright in that book. In contrast, sharing the single copy of the book you purchased with a friend would be well within your rights under copyright law.

---

81 See 17 U.S.C. § 1201(a).

82 For instance, until recently, Amazon’s Kindle supported only a proprietary AZW format instead of an open ePub format that lets consumers obtain an ebook from one company or organization and read it on a different company’s ereader. Cesar Cadenas, *Amazon Kindle Will Finally Support EPUB Books – But There’s a Catch*, TechRadar (May 3, 2022), <https://www.techradar.com/news/amazon-kindle-will-finally-support-epub-books-but-theres-a-catch>

83 This concern was also discussed by the senior VP of Grand Central Publishing in the Internet Archive Controlled Digital Lending litigation: “[E]books also carry special risks that can only be tamed by imposing limitations on how they are used. The downside of the replicability and portability of ebook files, for rightsholders at least, is that a single unprotected ebook file can satisfy a vast swathe of global demand for a title if it is allowed to proliferate.” Decl. of Ben Sevier 15–16, ECF No. 92. Penguin Random House’s senior VP for Online and Digital Sales made a similar assertion: “Anyone in possession of an unprotected digital file containing the text of a book can (for all intents and purposes) replicate it infinitely, instantaneously and at practically no cost.” Decl. of Jeffrey Weber 18, ECF No. 95.

84 “[limitations on how ebooks are used] also ensures that the municipal funding each library spends on ebooks is actually dedicated to patrons in that local community.” Decl. of Chantal Restivo-Alessi 11, ECF No. 94.

Although the DRM benefits both publishers and platforms by reducing purchaser control over ebooks, participants suggest that the use of DRM is driven primarily by publisher demands. None of them suggested that DRM was being driven solely by the platforms. Applying DRM is a major condition for publishers to work with ebook platforms at all.

“From a retailer perspective, we still believe it’s the publisher’s problem and not so much ours,” a digital director at an ebook platform company said. “If the publisher says, I will not give you these books unless you get with the times [and implement DRM], let’s say the industry is going in a certain direction. So then we would have to do software updates to support that. But we as a retailer are not driving that decision. We would just support where it goes.”

While DRM may be driven by publisher demands, it benefits platforms as well. Several participants agree that customers get hooked into ebook platforms and say the inability to easily transfer ebook files between different platforms, particularly with Amazon, makes it difficult for customers to switch between platforms, leading to a further concentration of market share among a few dominant players.

“With ebook, there’s Digital Rights Management on ebooks so you just don’t have the cross-pollination across platforms that one would hope for,” one participant said. “If you buy from Kindle, you cannot play it on a NOOK, you cannot play it on a Kobo device, so when you’re in the ecosystem of a company, that’s pretty much where you stay.”

Participants’ mixed views on the purpose of DRM tracked responses reported by industry analyst Mike Shatzkin over a decade ago. In an informal survey of “[n]ine very high-level executives in seven different top dozen publishing houses, plus four literary agents with extremely powerful client lists,” Shatzkin reported that “[e]ven of the 13 agreed with me that DRM is necessary to protect sales. Ten of the 13 agreed with me that DRM is not an effective deterrent to piracy. And 12 of the 13 agreed with me that DRM’s main benefit is to prevent casual sharing!”<sup>85</sup> Respondents then, and study participants today, understood that DRM’s practical effect was preventing the type of casual sharing between friends that is allowed under the First Sale Doctrine, and that DRM is not an effective anti-piracy measure.

One participant, who is an executive at an ebook platform company, explained how DRM is meant to limit the legal transferability of ownership, and not just prevent illegal piracy.

“Nobody wanted there to be a secondary market for ebooks, and I think mostly because nobody could imagine how you would control it or because you’re dealing with perfect replicas, like how do you ever keep that in the box?” the participant explained. “So the focus was on maintaining the perimeter through DRM, rather than a focus on ownership structure that could just prove to be impractical to try and enforce.”

---

85 Mike Shatzkin, *What the Powers-That-Be Think About DRM, and an Explanation of the Cloud*, The Idea Logical Company (Jan. 16, 2011), <https://www.idealog.com/blog/what-the-powers-that-be-think-about-drm-and-an-explanation-of-the-cloud/>

**The DRM placed on ebooks, when combined with ebook licensing schemes' limitations on first sale uses, can collapse the distinct activities of "lending" and "piracy" into a single violation of the license.**



The DRM placed on ebooks, when combined with ebook licensing schemes' limitations on first sale uses, can collapse the distinct activities of "lending" and "piracy" into a single violation of the license. In doing so, the DRM, and licensing limitations, can use the pretext of preventing an illegal activity (large-scale copyright infringement) to restrict a legal activity that has long been disliked by some publishers (loaning a copy of a book to a friend).

One participant at a Big Five publisher explained it like this: "We want customers to be able to read the book, to keep the book forever. What we don't want [is] to have someone take that ebook file and then share it with 10,000 other friends for free."

During the Internet Archive Controlled Digital Lending proceeding, the senior VP for Online and Digital Sales at Penguin Random House also conflated the ability to lend an ebook to a friend with worldwide piracy: "If the ebook could be shared among devices belonging to multiple users, the cost would be significantly higher to reflect the broader distribution. And the price to purchase an ebook file, without restrictions on its further use or distribution, would be utterly unaffordable as a practical matter given that it could potentially satisfy global ebook demand for that title."<sup>86</sup>

Copyright law prohibits a single user from sharing a single ebook with 10,000 friends. DRM does as well. However, only DRM prohibits a user from sharing a single ebook with a single friend.

The concern that individuals lending books might undermine sales was explicitly described by other participants as well. "So I think it's a conflicted sense around whether ebooks are good for publishing and good for the industry and whether or not they're kind of providing the basis for either undermining sales because of consumer sales by making it easier to lend and sometimes undermining sales because of piracy," said the participant from the trade book association.


As another participant in the digital side of one of the Big Five publishing houses explained, "[publishers] license because they want to make sure there is an artificial lifespan on the file, and an artificial cap on the sharing of that file."

Many participants appeared to believe that DRM does effectively prevent unauthorized copies of ebooks from appearing online. A former executive at a Big Five publishing house explained that "If there weren't enough protections, and the plain fact of the matter is, you know, 50 or 60% of the books out there were pirated, readily available for free, essentially, copyright would be gone."

---

<sup>86</sup> Decl. of Jeffrey Weber 20, ECF No. 95.

**Whatever other benefits DRM may bring to publishers, publisher insistence on applying DRM to ebooks has also undermined the development of a more-competitive platform landscape.**



Another participant in a director role at a platform company echoed this sentiment: “Once a book is out there in the world, there’s nothing to stop it from going viral. Especially, certainly, you’re never going to see that happen with a John Grisham or James Patterson [book]. ... There’s no way that they would make that DRM free.”

Both of these doomsday descriptions fail to account for today’s reality in which the majority of DRM-protected ebooks are already illicitly available online.

Other industry players appear to have less confidence in DRM as an anti-piracy tool. In January 2023, International Publishers Association (IPA) President Karine Pansa claimed that the fear of unauthorized distribution of ebooks was keeping entire regions out of ebooks despite the existence of DRM: “In Africa and the Arab world, we hear publishers saying that they won’t shift to digital because the risk of piracy is too great and their copyright enforcement regimes want to protect them.”<sup>87</sup>

Regardless of its actual effectiveness as an anti-piracy measure, DRM has effectively reduced readers’ freedom to own the books they pay for.

Finally, DRM creates a somewhat striking dynamic between publishers and the platform companies. Whatever other benefits DRM may bring to publishers, publisher insistence on applying DRM to ebooks has also undermined the development of a more-competitive platform landscape. It is notable that even though there are clear benefits to platforms when DRM is used to lock purchasers into specific walled gardens, the industry consensus is that the use of DRM is driven by the publishers themselves. Publishers’ decades-long focus on eliminating the secondary market may have led them to insist on technology that locked them into a handful of platforms.

## Platform Companies Use Walled Garden Infrastructure to Make Purchaser and Publisher Lock-In Profitable

Ebook platforms lock in publishers and purchasers by building digital walled gardens, while simultaneously collecting data from both parties. Ebook platforms lock in publishers with their distribution agreements, and they lock in users by allowing people to see ebooks only on the platform where they were originally

---

<sup>87</sup> Pansa, Karine. 2023. *A New Beginning: IPA @ Digital Book World*. Recorded at Sheraton New York Times Square Hotel, Jan. 16, 2023.

purchased (and prohibiting downloading, sharing, and other off-platform access).<sup>88</sup> This lock-in is in the interest of platforms, and creates some benefits for publishers. However, when it comes to walled gardens, the long-term interests of publishers and platforms may not be fully aligned.

Platforms leverage license agreements that prevent purchasers from taking their books outside the digital garden walls. This makes platforms sticky to purchasers. Once a purchaser has one ebook on a platform, they are likely to return for a second, third, and so-on. That allows platforms to convert their walled garden structure into increased ebook sales.

Ebook platform lock-in has specific impacts on libraries. Some platform schemes require libraries to pay for bundles of publications instead of picking out their own books. Before digital content was provided on internet-connected platforms, libraries had robust “collection development” teams tasked with reading through book reviews, advertisements, and other recommendations to curate their library collections title-by-title. That old model has been replaced by the modern digital platform model, in which a continuous stream of “content” is available on every library visitors’ “feed.”<sup>89</sup> In public libraries, these content collections can contain titles that are downright offensive or incorrect – in 2022, some librarians noticed that Hoopla, an OverDrive competitor, was platforming Holocaust denial propaganda and books containing incorrect information about reproductive healthcare, LGBTQIA issues, and other similarly problematic works.<sup>90</sup>

As one collections development librarian at an academic library told us, this purchasers’ mindset discourages libraries from sharing digital resources and developing unique collections: “Every single library that I know of was just throwing money at the problem [during the pandemic]. Like, fine, we’ll get all of the Project Muse ebooks. Fine, we’ll get all of the JSTOR ebooks. And so we just funneled money into these publishers, but we couldn’t actually be like, okay, [my library’s] buying all of Johns Hopkins so that Yale can buy all of JSTOR and we’ll just lend. Like we couldn’t do that. We all just bought the same stuff and that’s terrible from a collection development standpoint.”

This study participant believes academic libraries are united in thinking this approach isn’t working, to the point that some university librarians are turning their collection focus back to purchasing print books: “Because [print books are] a for-sure thing, like there’s a physical object. We can lend that physical object,” the study participant added.

---

88 SPARC, the Scholarly Publishing and Academic Resources Coalition, describes how library resource platforms lock in its report about library services platforms, describing switching platforms as expensive and time-consuming, and comparing the process to “a kind of open heart surgery for the library.” Letter Opposing the Merger Between Clarivate PLC and ProQuest LLC 10 (Oct. 22, 2021), <https://sparcopen.org/wp-content/uploads/2021/10/SPARC-FTC-Letter-in-Opposition-to-the-Clarivate-ProQuest-Merger.pdf>

89 Theodore C. Bergstrom et al., Evaluating Big Deal Journal Bundles, 111 PNAS 9425 (2014).

90 Claire Woodcock, *Ebook Services are Bringing Unbinged Conspiracy Books into Public Libraries*, VICE (Apr. 10, 2022), <https://www.vice.com/en/article/93b7je/ebook-services-are-bringing-unbinged-conspiracy-books-into-public-libraries>

Libraries may choose not to participate in the ebook market because of what they perceive as restrictive, and sometimes seemingly exploitative, terms of use. The platform providers tack on digital bells and whistles to the items they sell, like product-related metrics, customer ratings, and algorithm-driven “see more like this” options. These data analytics and algorithmic features add value to the copyrighted material by giving consumers extra information about the content they’re viewing, but they are also tools for publishers to assert more control over the materials they provide to consumers. By embedding add-ons into items that used to be sold outright, publishers can claim that the raw material of their publications are inseparable from the platforms and metadata they’re attached to.<sup>91</sup>

As discussed in more detail below, these features are also tools for personal data collection and use. Walled garden platforms are uniquely poised to provide their owners (or those with access to their data) a “God’s-eye view” of their consumers’ habits, preferences, and decision-making tendencies. Publishers can use customer ratings and other consumer feedback to intuit what particular users will want, based on their data, and predict what they will buy in the future, and to decide what books to publish and what prices to charge.

## Publishing Platforms Also Lock In Publishers

Platforms don’t just lock in readers. They also lock in publishers.

Market-dominating platforms like Amazon and OverDrive have emerged as obligatory content middlemen that publishers must cooperate with in order to have their books seen by most consumers. According to one estimate, Amazon leads the US ebook market with around 81% of the market share in 2020.<sup>92</sup> OverDrive claims at least 90% of public libraries in the US and Canada as clients.<sup>93</sup>

---

91 Legal information publishers, like Westlaw and Lexis, have been making claims like this for decades, alleging that everything from their case law headnotes, which summarize different legal holdings in court opinions, to “star pagination” that notes where page breaks are in original legal text, give them copyright over otherwise-public, un-copyright-able texts.

92 *EBook Market Share: Amazon, Apple and Kobo*, Booksliced, <https://booksliced.com/books/the-exact-ebook-market-shares-of-the-major-players-in-the-industry-are-rather-difficult-to-come-by-but-here-are-some-estimates> (last visited Apr. 1, 2023).

93 Marshall Breeding reports that OverDrive supplies its ebook platform to over 95% of public libraries in the US and Canada. Marshall Breeding, *OverDrive to Change Ownership*, Smart Libraries Newsletter (Feb. 2020), <https://librarytechnology.org/document/25069>; A 2021 *Washington Post* report said that almost 90% of US libraries use OverDrive’s Libby app to connect patrons to ebooks and audiobooks. Geoffrey A. Fowler, *Want to borrow that e-book from the library? Sorry, Amazon won’t let you*, *Wash Post* (Mar. 10, 2021), <https://www.washingtonpost.com/technology/2021/03/10/amazon-library-ebook-monopoly>



**Walled garden platforms are uniquely poised to provide their owners (or those with access to their data) a “God’s-eye view” of their consumers’ habits, preferences, and decision-making tendencies.**



Participants regularly described how the limits of Amazon’s market-dominating Kindle, in particular, undermined economic incentives to innovate.

“We would want to create better versions of children’s books, you know, something that has, like, better picture quality, better audio, whatever. But Amazon’s devices don’t support audio. So immediately you just can’t have a read-along children’s book,” explained one participant from a Big Five publisher.

Another participant from a trade association echoed that sentiment: “That’s one of the reasons why ebooks suck. They don’t really take advantage of the medium at all [and] the least common denominator is Kindle. It’s kind of a black-and-white provisioning – Amazon doesn’t even fully support accessibility for print-disabled audiences. We could be doing a lot more with digital formats. But we’re not because the primary distribution mechanism, Amazon, you know, 80% of the market, doesn’t really care about innovation in the ebook market. And as a result, we kind of puddle along.”

In part, this dependence on third-party platforms is because many publishers prefer to outsource the process of creating and supporting ebook files to third parties. A director of an ebook distribution platform explains that ebook distribution is “just not [the publishers’] focus. Like, they’d have to create a whole structure like an ecomm [an ecommerce platform]. And some of them will have it, like [Penguin Random House] has it but, you know, an e-comm structure that gets supported and iterates... it’s a business, it’s a massive undertaking, quite honestly. And if their focus is to bring content to the world, and make some money off of it, it’s a distraction.”

Another participant working at a Big Five publisher agreed: “It just wasn’t feasible, like the amount of time and effort put into developing and maintaining the site or a space where customers can redeem their files from, really did not represent the actual sales that we were getting through those stores.”

Amazon, and other popular consumer platforms for ebooks, use their power in the ebook market to directly contradict publisher preferences. They offer publishers’ books for prices below their retail price, making it harder for publishers to recoup their publishing costs.<sup>94</sup> They also leverage the consumer data they collect on their platforms to prevent competing platforms from taking

94 *Id.* (“In July, the Authors Guild organized a lawsuit against a notorious Ukraine-based network of piracy sites called Kiss Library. The named plaintiffs in the suit are 12 Authors Guild members, as well as Amazon Publishing and Penguin Random House, whose generous financial support made the lawsuit possible. Kiss Library, which operates Kissly.net, Libly.net, Cheap-Library.com, and other websites, illegally sells pirated ebooks at very cheap prices.”)



hold in the market,<sup>95</sup> and they have even created their own self-publishing services that route authors away from the traditional publishers, setting up a new, rival publishing industry.<sup>96</sup>

“There’s a lot of issues tied up with the publishing industry’s relationship with Amazon, and it’s a difficult subject to speak on,” a journalist on the book industry beat said. “I’m not gonna guarantee that someone’s not going to burst through my door here and put a bag over my head and throw me in the back of the van [for agreeing to participate in this research study].”

This ownership of the ebook customer base can also give platforms like Amazon control over which books readers discover in the first place. “The ways that you would make a decision about how to buy those books feel like they’re much more controlled by the apps and what the apps recommend to you, as opposed to like, this author recommends it and it got this review,” one study participant, who wishes to remain anonymous, said.

One participant with a market research group pointed out that data about consumer behavior tends to stay with the platforms and not be shared with the publishers.

“Unlike Elsevier, they don’t have a good view into [data]” the participant said. “The platforms that do own themselves... Those guys can and do use data like that to optimize their catalogs. But the publishers themselves don’t have visibility into that. And the main players like Kindle will never share it. So [publishers] are kind of at a disadvantage in that respect.”

Platforms can also disrupt existing publisher-purchaser relationships. In 2020, the Association of American Publishers, the Authors Guild, and the American Booksellers Association wrote to the House Antitrust Subcommittee to urge the government to step in and provide more oversight to protect the publishing industry from Amazon.<sup>97</sup> The publisher organizations blamed Amazon for not competing on a “level playing field when it comes to book distribution,” and claims that the ebook giant “owns and manipulates the playing field, leveraging practices from across its platform that appear to be well outside of fair and transparent competition.” And because the company is the dominant platform for so much of the ebook market, publishers have no choice but to try to make the relationship work, even though Amazon exploits them.

---

95 Jim Milliott, Publishing Leaders Issuing Warning over Amazon’s Market Power, *Publishers Weekly* (Aug. 18, 2020), <https://www.publishersweekly.com/pw/by-topic/industry-news/bookselling/article/84119-publishing-leaders-issuing-warning-over-amazon-s-market-power.html> [hereinafter Milliott, Publishing Leaders Issuing Warning over Amazon’s Market Power]

96 Jeffrey A. Trachtenberg, ‘They Own the System’: Amazon Rewrites Book Industry by Marching Into Publishing, *WALL ST. J.* (Jan. 16, 2019), <https://www.wsj.com/articles/they-own-the-system-amazon-rewrites-book-industry-by-turning-into-a-publisher-11547655267>

97 Joint Publishing Industry Letter to Chairman Cicilline, *supra* note 16; Milliott, Publishing Leaders Issuing Warning over Amazon’s Market Power, *supra* note 69 (discussing joint publishing industry letter).

## Publishers and Purchasers Explore Creating Their Own Platforms

Both publishers and purchasers have attempted to sidestep some of the challenges presented by independent platforms by creating their own.

There is some evidence that at least some publishers see the value in developing their own ebook distribution platforms. Both Simon & Schuster and Penguin Random House pride themselves on developing robust ebook creation and distribution pipelines. According to one participant familiar with the operation, Simon & Schuster is beginning to see smaller publishers leave relationships with OverDrive and enter into distribution agreements with them instead. Direct ownership over its own ebook distribution could help Simon & Schuster avoid problems that arise when its interests do not align with the interests of platforms.

There are also examples of institutional purchasers working to create their own platforms. New York Public Library (NYPL) is one of the only public library systems in the country to launch its own library lending app, which one NYPL employee said became an important initiative for reinstating contact with patrons.

“Our patrons were seeing content, content was being promoted to our patrons, that is not stuff that we would necessarily have promoted,” one NYPL employee familiar with the project said. “We were getting concerned about what that relationship looks like. Sometimes we would hear patrons talk about how, you know, they check things out from Libby [an app from OverDrive]. Well, they actually check things out from NYPL, through Libby. We were starting to lose some of that connection. [NYPL developed its own platform because] we wanted to have more control over the relationship with the patrons but also with the publishers.”

This fear of disintermediation was echoed by the publishers in their complaint against the Internet Archive’s Controlled Digital Lending program. “[Internet Archive’s] website undercuts public libraries by disintermediating them.”<sup>98</sup>

## Publishing Platforms Offer New Benefits to Readers and Publishers

While the platforms impose limitations on both purchasers and publishers, they also provide benefits. Ebook platforms give readers a reprieve from the performative aspects of reading a physical book in public. It’s no wonder romance fiction is one of the top ebook genres, according to study participants familiar with general purchaser habits and trends.

---

98 Compl. 46, ECF No. 1.

Ebooks also have a functional appeal for some reader demographics, a director at a leading market research company explained: “Older readers tend to really like them because they can do things like have larger font sizes, they can adjust the font, they’re easier to read for people who have or are starting to have issues with their vision.”

A book industry consultant also talked about the distinction between ebooks with plain text versus ebooks with pictures and features like hyperlinks: “Ebooks have functional capabilities that print books do not have.”

Platforms offer services that keep customers coming back. For writer Louryn Strampe, Kindle’s built-in dictionary was particularly helpful when rereading a book like *Gravity’s Rainbow*.<sup>99</sup> Other readers value knowing that an ebook listed by a platform is always available.

“They’re never out of stock,” the consultant said. “They’re never out of print. You can always deliver the next one.”

The platforms also aggregate customers and titles, increasing convenience for consumers and drawing them to books they may not ordinarily see in the course of their physical shopping errands. On digital platforms, consumers can access a particular ebook, song, or video while also interacting with elements of the online store. Consumers can shop online, browsing algorithmically curated digital collections instead of merely getting a singular, pre-selected item; personalize the items they’ve bought by creating online annotations and filing systems to hold their favorite products; leave comments and have conversations with other shoppers; and rate and rank the items they’re perusing.

While some platforms invest deeply in algorithmic recommendation and curation, other platforms intentionally avoid relying on these tools.

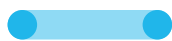
“Because we’re a bookseller, I think we have good curation,” one director at an ebook platform company said. “The stores are already making lots and lots of decisions every day on what they want to get behind and what’s important to the company and the idea of elevating certain writers that we think should be elevated. ... Ebook, in a lot of ways, just follows what the larger goals are.”

At the same time, it is important to disentangle the benefits of platforms as a concept from the version of platforms that exist today. There is no reason that easy digital access to books must be paired with proprietary file formats that prevent customers who purchase books from bringing them elsewhere, or transferring them to someone else. There is also no reason that digital access to books must involve data extraction from consumers for marketing and other commercial purposes. Almost all platform features could also function in a world where customers owned their ebooks and were free to bring them to (or remove them from) platforms as they saw fit, without having to worry about the privacy and use of their ebook data. Similarly, there are many examples of digital goods that are sold, not licensed, online.<sup>100</sup>

99 Louryn Strampe, *Ebooks Made Me Fall Back in Love With Reading*, *WIRED* (Nov. 21, 2022), <https://www.wired.com/story/kindle-ebook-digital-reading-rave/>

100 For example, the music website Bandcamp makes no attempt to frame digital purchases of songs or albums as licenses.

# Monetizing Data and Tracking Users Provide Additional Incentives for Publishers and Platforms to Retain Control



Platforms don't just let publishers license instead of sell ebooks or lock consumers into particular walled gardens. They also create entirely new money-making opportunities in personal data and data analytics markets. While both publishers and platforms are attempting to use and monetize this data, platforms appear to be more aggressively invested in its use.

Locking readers into platforms generates data that was formerly impossible to collect. Today, platforms are able to gather location data on users, both in terms of what page they're on and where they're using their device. This is in stark contrast to what publishers or bookstores can track when a customer walks out the door of a bookstore with a paper book in hand. That book sale transaction has ended at the point of sale, making it impossible to know if they read the book at home, on a plane, or in a hospital. It's impossible to know whether a reader lingers on a particular page. It's impossible to know if a reader recommends the book to their friends. All of this data can now be harvested from ebook platforms.<sup>101</sup>

---

<sup>101</sup> Another cottage industry has developed around managing and running analytics on publisher data. *See, e.g.*, DoubleVerify, <https://pub.doubleverify.com/> (last visited Apr. 1, 2023); Hurix, <https://www.hurix.com/technology-services> (last visited Apr. 1, 2023).

That data can be especially valuable when combined with other data and sold through data brokers. Data brokers specialize in collecting personal data or data about companies and selling or licensing that information to third parties for any number of reasons. Personal data collection and sale is a multibillion-dollar business. Data brokering, sharing, and sale to advertisers, companies, and even public institutions is a ballooning nascent business. Ebook platforms give publishers the opportunity to participate in the personal data industry. Many platforms are already there.

Ebook platforms require customers to provide personal information in order to access ebooks. This information can include your name, address, digital payment information, and whatever other personal data points the platform owner wants. Once you start using the platform, your every move on the platform becomes a data point that could be collected. While it's not clear what types of data various platforms collect, people have discovered that Amazon tracks every time a user taps their Kindle ebook reader,<sup>102</sup> and companies like Gale sell data analytics that tell libraries what decisions to make by “blending ebook and audiobook vendor data” with “segmentation and lifestyle data.” Gale’s ebook reader data is specific enough to mark readers’ locations on a map and provide information including their estimated household income and how many children they have.<sup>103</sup>

Consumer data is valuable, both to ebook platforms, and as a separate saleable item that ebook platforms and publishers can barter and sell with others. As one journalist with experience covering the topic put it: ebooks really give Amazon a wealth of customer data. “It really helps build out a customer profile, and I think that’s one of the major things that Amazon has going for it that publishers don’t, and that libraries don’t, because Amazon owns the customer. They own the customer profile. They own the data, and that’s just a huge part of how the digital environment functions.”

Platform services are a conduit for consumer data, allowing companies to track their consumers’ choices. Ebook platforms can track consumers, not just at the point of sale and beyond, but before sales occur on the platform. The companies can lead consumers to products based on their personal online habits and prior choices.

---

102 Chaim Gartenberg, *Why Amazon is tracking every time you tap your Kindle*, THE VERGE (Jan. 31, 2020), <https://www.theverge.com/2020/1/31/21117217/amazon-kindle-tracking-page-turn-taps-e-reader-privacy-policy-security-whispersync>

103 *Gale Analytics: Digital Collections*, Gale, <https://www.gale.com/c/gale-analytics-ebook-and-audio-book-insights> (last visited Apr. 1, 2023).

It's no surprise that ebook platforms are following most other industries and infusing their ebook readers and ebook selection platforms with internet-connected data collectors. While it's not clear how or if trade publishers will use this type of data, academic publishers are learning how to leverage readers' data to manipulate the pricing of their ebooks, depending on who is using them, and by generating metrics that predict how impactful various publications will be.<sup>104</sup>

Ostensibly, when an ebook provider can see exactly what every one of its customers does with each of its products and services, it can use that data to understand what consumers want, even before the consumers themselves know what they want.<sup>105</sup> This data can also distort what books are published, replacing human editorial evaluation of new works with data-driven decisions that act to replicate what has succeeded in the past.<sup>106</sup> Today, "Publishers can know that you raced through a novel to the end, or that you abandoned it after 20 pages. They can know where and when you're reading. On some reading sites and apps, users sign in with their Facebook accounts, opening up more personal data."<sup>107</sup> But it's not always the case that the companies that run the platforms are sharing the data they collect with publishers. For instance, companies including Amazon and Apple are not sharing user data with publishers.<sup>108</sup>

Other publishing industries, including the academic publishing industry and the legal information industry, have figured out how to use readers' data to build new, profitable data "insights" to sell to industry consumers. The publishers have gleaned data from their users on their journal and legal information platforms, and used their consumers' data to develop

---

104 In 2022, an academic publisher, Wiley, pulled over 1,000 of its books from a platform called ProQuest in order to charge more for those popular titles, which Wiley probably assessed by user data showing who was accessing what books, and how frequently consumers were accessing them, etc. Gary Price, Wiley Plans to Restore Access to Approximately 1,380 E-Book Titles Removed from ProQuest Academic Complete Collection, *Library Journal* (Oct. 5, 2022), <https://www.infodocket.com/2022/10/05/wiley-plans-to-restore-access-to-approximately-1380-e-book-titles-removed-from-proquest-academic-complete-collection-titles-will-return-as-soon-as-possible>; Academic publishers are also known for using their consumers' data to create "impact factors" that judge how important academic publications are based on factors such as how often people click on them. Sandy DeGroot, *Measuring Your Impact: Impact Factor, Citation Analysis, and other Metrics: Other Metrics/Altmetrics*, University of Illinois Chicago Library, <https://researchguides.uic.edu/c.php?g=252299&p=1683203> (Feb. 9, 2023, 2:02 PM).

105 Thomas H. Davenport et al., Know What Your Customers Want Before They Do, *Harv. Business Rev.* (Dec. 2011), <https://hbr.org/2011/12/know-what-your-customers-want-before-they-do>

106 Susanne Althoff, Algorithms Could Save Book Publishing – But Ruin Novels, *WIRED* (Sep. 16, 2016 7:00 AM), <https://www.wired.com/2016/09/bestseller-code/> (discussing the book *The Bestseller Code*, about an algorithm that can predict whether a book will be a bestseller with 80% accuracy).

107 *Id.*

108 *Id.*

“academic metrics” and “legal insights” data products.<sup>109</sup> These products predict which scholars’ projects will be the most lucrative and tell lawyers how particular judges are likely to react to their arguments based on their troves of personal data and other copyrighted informational troves (news, academic journal articles, court opinions, etc.). These other publishers have figured out that if you combine information from the materials you provide on your platforms with personal data from the consumers who use those platforms, you can derive brand-new information products to sell.

But these types of products are not yet prevalent on trade ebook platforms, nor are they guaranteed to appear on those platforms in the future. “I think academic publishing, the Elseviers of the world, are not easy models for trade publishers to look at but they’re useful ones,” one participant from a book trade association that deals with standards and policy said. “We think about these things a lot, even though we have a trade publishing bias in our membership. The reality is that we’re always thinking across the industry, we have to. And I think that it will come. My job is to hopefully have it come at a time and with capabilities that support a transition to a different way of working because I don’t want it to disrupt book publishing, I want to help book publishing get better and bigger.”

Regardless of how ebook publishers and platforms are using their licensees’ and platform visitors’ data, they understand the potential value of that data and can see how academic publishing has created an entirely new data analytics market (selling author metrics, sales predictions, etc.) that is a multibillion-dollar industry.<sup>110</sup> Ebooks platforms will likely, eventually, learn how to harness the profit opportunities of “big data.” It is less clear if the publishers themselves will follow suit.

## Publishers Are Working to Use Reader Data

Most ebook platforms are already collecting myriad customer data. It is harder to understand exactly how publishers are using the data they have access to, which may only be a portion of the data collected by platforms. In our current data privacy law regime, nothing requires the

---

109 Lamdan, *supra* note 3. Cody Hanson, *User Tracking on Academic Publishing Platforms*, Cody Hanson, <https://www.codyh.com/writing/tracking.html> (last visited Apr. 1, 2023); Todd A. Carpenter, *Elsevier Has Deployed an End-user Tracking Tool for Security*. Should Users Be Concerned About Their Privacy, Scholarly Kitchen (Oct. 13, 2020), <https://scholarlykitchen.sspnet.org/2020/10/13/elsevier-has-deployed-an-end-user-tracking-tool-for-security/>

110 Jeff Pooley, *Surveillance Publishing*, *The Journal of Electronic Publishing*. (2022) 25(1). doi: <https://doi.org/10.3998/jep.1874> (discussing the multi-billion dollar deals to acquire bibliometric data analytics tools among academic publishers and academic analytics companies).



companies collecting our data to tell us what data they are collecting or what they are doing with that data. The platforms are “closed” and opaque,<sup>111</sup> and most of the data they generate are “proprietary and purposefully locked away.”<sup>112</sup>

We do know that Penguin Random House, one of the Big Five trade book publishers, is building out its data analytics departments. The company is trying to harness artificial intelligence to extend the reach of its ebooks<sup>113</sup> and trying to “develop their competitive advantage by transforming their data solutions and leveraging their vast data assets” by partnering with consumer data-leveraging companies like Snowflake to better use data from Amazon and other platforms.<sup>114</sup> One publisher at the Big Five who uses these tools says they aren’t that powerful yet: “We can use an AI generation tool to add keywords to the metadata of a book, but I don’t think it makes that much difference for search on online platforms.”

Following in the footsteps of legal and academic publishers, Penguin Random House is developing algorithms and other techniques to benefit from the personal data collected from its digital services. The book company is actively seeking employees who can “[build] upon the latest results from the academic research community.”<sup>115</sup>

Trade publishers have been slower to integrate data analytics into their own publishing decisions, and to find ways to monetize data analytics by selling them to interested customers (book vendors, writing programs, entertainment companies, etc.). When publishers were focused on selling paper books, information about consumers’ online behavior was more relevant to the digital retail platforms that sold their books than the publishers themselves.<sup>116</sup> Today, “... taking ownership of reader data appears to be increasingly important to publishers.”<sup>117</sup> Data analytics are being developed to do more useful work for publishers, making tools that can help target readers and predict publishing market trends.

---

111 See Michael R. Baye et al., *Searching for Physical and Digital Media: The Evolution of Platforms for Finding Books*, Nat’l Bureau of Econ. Rsch. 7 (Oct. 2013), <https://www.nber.org/papers/w19519>

112 Melanie Walsh, *Where Is all the Book Data?*, Public Books (Oct. 4, 2022), <https://www.publicbooks.org/where-is-all-the-book-data/>

113 Phil Stamper-Halpin, *How PRH Uses Artificial Intelligence (AI) to Extend Your Book’s Reach*, Penguin Random House (May 2019), <https://authornews.penguinrandomhouse.com/how-prh-uses-artificial-intelligence-ai-to-extend-your-books-reach-2/>

114 *Penguin Random House: Plotting a story of data maturity with Snowflake*, Snowflake (Apr. 8, 2022), <https://www.snowflake.com/blog/penguin-random-house-plotting-a-story-of-data-maturity-with-snowflake/>

115 Job Posting for Machine Learning Manager, Penguin Random House, <https://perma.cc/MJ9T-XCKX>

116 Sybil Nolan and Alexandra Dane, A sharper conversation: book publishers’ use of social media marketing in the age of the algorithm, *Media Int’l Austl.* 168, 153–166 (2018), <https://journals.sagepub.com/doi/full/10.1177/1329878X18783008>

117 *Id.*

**It seems trade publishers haven't consistently been able to figure out the best way to harness data analytics and are not yet using them as primary determinants in publishing decision-making, nor are they reselling them to the extent that academic publishers are.**



One 2018 study of publishers reflected their moves to utilize personal data analytics in their marketing strategies. All of the big publishers had started using website analytics and social media metrics on top of the traditional “point-of-sale” data collected by Nielsen BookScan.<sup>118</sup> A 2016 report focused on how Penguin Random House UK was using data analytics to “paint a detailed picture” of their readers, and to sort them into categories to “connect the right readers with the right books” based on their “attitudes, behaviors, and motivations.”<sup>119</sup> This data analytics work allows publishers “to sidestep the time-consuming, costly, and often haphazard work of identifying or creating a unique audience for each and every title in their catalogs.” One independent publisher said, however, that “it’s really hard to take shortcuts here” because each book and each author is so different. The publishers who can afford robust data analytics tools use them to “inform how [they] acquire books,” pick their publications, and market them.

Much of the data analytics used by publishers is still not at the level used by retail platforms like Amazon. It seems trade publishers haven’t consistently been able to figure out the best way to harness data analytics and are not yet using them as primary determinants in publishing decision-making, nor are they reselling them to the extent that academic publishers are. This is especially true for smaller publishers. Large publishers are investing more in data analytics as part of their marketing strategies.<sup>120</sup>

In fact, it seems that it is still relatively difficult to obtain current, accurate data on book sales and other publishing data.<sup>121</sup> It also seems that, currently, data does not necessarily guide many of the decisions within the industry. In discussing the use of data with participants who work for and with publishers, a distinction emerged between marketing and production data on one hand, and manuscript evaluation on the other.

“This is a business that’s 500 years old,” said one participant from a trade association. “You know, even in the US and the current model, it’s close to 100 years old. And the reason is that we don’t use data. I mean, we just don’t really think about it. People, you kind of do shortcuts and shorthands.”

---

118 *Id.*

119 *Id.*

120 *Id.*

121 Melanie Walsh, *Where is All the Book Data?*, Public Books (Oct. 4, 2022), <https://www.publicbooks.org/where-is-all-the-book-data/>

Manuscript evaluation was regularly described as something that existed outside of the world of data. “No one wants to believe that a machine can read a book and tell you anything about the book,” the same participant said. “This is what editors do. And it can only be done in an analog way reading one manuscript at a time. And you know a good book when you see it. And that’s the only way to choose a book.”

This sentiment was echoed by one former Big Five executive: “Their editors, senior editors, executive editors, publishers, and it’s all sort of intuitive, right? So it’s a business based on intuition.”

Instead of being used at the core of publishing and editing choices, data is being used for manufacturing and marketing decisions. Online tracking allows publishers to trace the journey from ebook to newsletter signups. “We also have links where it’s like, sign up for our mailing list and newsletters. We also track that to see how many customers actually, you know, subscribe to our mailing lists for ebooks,” said one Big Five employee.

One former executive leader at a publishing house explained how this information actually shaped manufacturing choices: “But to me, it should be about making the back-end efficient, making good decisions based on real data as opposed to oh, I feel really good this morning so I think I’m going to spend more money. The enhancing ability to run a productive business, you bet. How many books should we print? Human intuition should not enter that. How many orders do we have? What’s the rate that we’re selling the book at? What are the channels of distribution? Here’s how many we’re going to print, here’s the cost of the book. Here’s how long it takes us to go back to press on this type of book. Here’s the cause of piles. Man, that’s a very complex decision. Actually, do we reprint this book right now? Sometimes [it’s] really easy. But when it’s a tough decision, it’s really tough. And you need lots of data.”

There are places where marketing and manuscript evaluation can overlap. One participant on the digital side of one of the Big Five described a system where branded cookbooks could be quickly created in response to industry trends. The publisher maintained a content management system with 10,000 recipes that could be repurposed and repackaged, “They would see like, you know, okay, this Stranger Things is trending or, you know, like, which, suddenly, a lot of like witchy products are trending. So let’s make a witchy cookbook.”<sup>122</sup>




---

122 “If you’re going to spend a lot of time trying to get your content in as many places as possible, then what are the ways in which you can get to the ideal reader to discover that content? How do you create marketing material? How do you communicate through social media?” This is where trade publishers are putting the greatest focus [...] Trying to use data to impact book creation, he added, is still a long way away.” Ellen Harvey, *How Cengage, HarperCollins, and Scribd Are Using Data Analytics*, *Book Business* (Apr. 29, 2015), <https://www.bookbusinessmag.com/post/how-engage-harpercollins-scribd-are-using-data-analytics/>

# Conclusions



All ebook stakeholders understand that licensing terms, DRM, and walled garden platform infrastructure place additional controls on how ebooks are used. They simply disagree as to whether or not those limitations improve or degrade the ebook economy, ecosystem, and experience.

This disagreement can be context specific. For example, a former Big Five executive's explanation for why he does not use Spotify for his music may also resonate with concerns that others have expressed about having to license ebooks: "I don't do Spotify and all that stuff. I do [individually purchased tracks through] Apple Music. Because I put it on my machine, and I do some of it through the cloud, then what I do is, every few years, I take a machine, and I take it offline, and I have my music. So I can always then put that back on my computer and have my music. Because you know, what happens if, I don't know, a weather event wipes out a significant amount of the server farms in the United States? Oh you're not gonna be able to get your stream of music or stream of anything else. And you see what Apple does with music is fantastic, right? They, in the new terms that you sign on in order to have your iTunes and all those songs you bought over the last 25 years. They have the right to swap out versions and they do it all the time. So suddenly your favorite acoustic version of that song that you love and you've been listening to for 20 years, you put it on and it's a different song. It's the same song, but it's a different performance. They just swap it out. So there is that issue of 'do I actually own it?'"

**The stakeholders in the world of ebooks are all working toward rational goals, whether it's growing a business, expanding access to knowledge, or simply wanting to read in private. Legal and technological changes have created a new market structure that tends to privilege the interests of publishers and platforms over those of purchasers and readers.**



The concern about swapping out versions of songs has direct parallels in ebooks. Amazon famously deleted copies of the novel *1984* from some users' devices after a licensing dispute.<sup>123</sup> More recently, the estate of Roald Dahl attempted to quietly revise the text of many of the author's books in new editions of physical books,<sup>124</sup> as well as retroactively changing the text of previously purchased digital versions.<sup>125</sup>

The stakeholders in the world of ebooks are all working toward rational goals, whether it's growing a business, expanding access to knowledge, or simply wanting to read in private. Legal and technological changes have created a new market structure that tends to privilege the interests of publishers and platforms over those of purchasers and readers. This change is especially stark when benchmarked against the physical book status quo. Fortunately, some straightforward interventions may help preserve the benefits of ebooks, protect against stakeholder concerns, and bring the relationship between publishers, platforms, and purchasers into historical balance.

## Regulators Should Address Competition Concerns and Barriers to Entry for the Ebook Marketplace

As this paper and other research shows, the shift to publisher-platform partnerships and the accompanying switch from sales to licensing have steadily gained both momentum and intensity in recent years. While there is no direct evidence to suggest explicit coordination, the increasing concentration in both the publishing and platform markets for ebooks, coupled with the massive simultaneous shift from sales to licensing, raises important questions about the nature and extent of any implicit or explicit coordination. For example, the near-universal refusal to offer individual or institutional consumers the option to legally own an ebook at any price raises the possibility that some agreement might be responsible. The absence of this option, on top of publisher-platform-reader lock-in, whereby access to books is prohibited on other platforms, could create substantial barriers to entry for new participants.

123 Brad Stone, Amazon Erases Orwell Books From Kindle, *NY Times* (July 17, 2009), <https://www.nytimes.com/2009/07/18/technology/companies/18amazon.html>

124 Ed Cumming et al., The Rewriting of Roald Dahl, *Telegraph* (Feb. 24, 2023, 4:20 PM), <https://www.telegraph.co.uk/news/2023/02/17/roald-dahl-books-rewritten-offensive-matilda-witches-twits/>

125 Ben Ellery and James Beal, Roald Dahl ebooks 'force censored versions on readers' despite backlash, *The Times* (Feb. 25, 2023), <https://www.thetimes.co.uk/article/roald-dahl-collection-books-changes-text-puffin-uk-2023-rm2622v10>

Two solutions can address these problems. One is to mandate that any ebook purchased should be delivered in a format that allows interoperability between ebook reading platforms. This would allow both purchasers and publishers to shift from platform to platform depending on which ones offer them the most competitive and preferred treatment, including new entrants to the market. The second solution is to mandate that vendors offer the option to “buy” ebooks (as opposed to rent/stream/license or more “temporal” transactions). This would require vendors to offer the ebook with the incidents of digital ownership, including the right to preserve the copy they buy, resell it, and be free of surveillance while they read it.<sup>126</sup>

## Economic Examinations of the Ebook Market Should Distinguish Between the Market for Digital Ownership of Ebooks and the Market for Platform-Controlled Licensing of Ebooks

At the same time that publishers have been investing more heavily in ebook platforms and less in traditional ebook sales, they are suing one library, the Internet Archive, for offering ebook loans under a system called Controlled Digital Lending (CDL).<sup>127</sup> CDL is a special system for library ebook lending that uses technology to follow a first sale framework. For example, libraries can lend out only one ebook copy at a time for each physical copy of the book they own.<sup>128</sup> This “owned-to-loaned” ratio simulates the economics of first sale for physical library books, where each physical book owned can be loaned to only one patron at a time. CDL is currently being implemented by a wide range of libraries in a wide range of ways.

The reasons behind the CDL lawsuit may be related to publishers’ ongoing belief that library loans cost them profits, and that shutting down digital library loans will recoup those losses. As noted in this paper, that belief was shaped by a physical book environment where readers were choosing between purchasing a book from a bookstore or borrowing a functionally identical copy of a book from the library.

However, the ebook marketplace looks very different from this physical book model. Bookstore ebook sales are a small and diminishing part of the market. Instead, licensed copies distributed via surveillance-driven platforms dominate the market, generating revenue from individual and bulk purchases, as well as subscriptions, advertising, and data analytics. The complexities

---

126 See Aaron Perzanowski and Chris Jay Hoofnagle, What We Buy When We Buy Now, 165 *U. Pa. L. Rev.* 315 (2017).

127 See *Hachette v. Internet Archive*, Elec. Frontier Found., <https://www EFF.org/cases/hachette-v-internet-archive> (last visited Apr. 1, 2023).

128 See David R. Hansen and Kyle K. Courtney, A White Paper On Controlled Digital Lending Of Library Books, *Controlled Digital Lending* (2018), <https://controlleddigitallending.org/whitepaper/>

**The complexities of these platform economics raise serious questions about how digital library lending, especially under a CDL approach, competes with current publisher business models.**



of these platform economics raise serious questions about how digital library lending, especially under a CDL approach, competes with current publisher business models. This ambiguity is particularly relevant to the Internet Archive lawsuit, as one of the four “fair use” factors at issue in the case is the effect of CDL on the market for the books being lent out. The greater the difference in the economic model of digital ownership versus platform-controlled licensing, the less market harm publishers can prove.

## Digital Ownership Should Be Normalized Instead of Publisher Refusal to Sell Ebooks

The deals that vendors strike with libraries on behalf of publishers often come with nondisclosure agreements libraries are expected to sign, waiving their right to talk about the different deals vendors are able to score with different libraries. This anti-competitive practice discourages librarians from inquiring about the deals vendors make with other libraries to determine whether or not they are getting a raw deal. Through our interviews with different participants, we learned that some academic libraries are opting to return back to purchasing physical books as a way to get around the constraints of ebook licenses.

Academic publishers, including MIT Press, are looking at preservation in terms of perpetual access, even though public libraries, who are not frequent purchasers of academic press books, are more likely to concede to a perpetual access license because the mission of public libraries is more preservation-oriented. One source with a larger metro public library system said that it viewed perpetual access as a form of preservation if proper backups are in place: “If we’re licensing ebooks and we get a perpetual license to those ebooks, there’s always the threat that the company that’s licensing it to us goes down and we lose that access. So we’ve built some preservation requirements into these things that say that you the publisher actually have to deposit copies into escrow, and they sit on this server that we can’t access until or unless the company goes bankrupt. At that point, the server unlocks and now we get access to all this content that we actually paid for. But now we get to preserve it because the company that licensed it to us originally went away.” Whether more libraries can negotiate these types of contractual agreements depends on how well resourced the staff at each system is, as well as how this would apply to their specific collection priorities and needs.



## Library Privacy Laws Should Be Updated to Preserve Intellectual Freedom on Ebook Platforms

Academic publishers are already using purchasers' personal data to assess researchers' impact and to do other data analytics. Right now, it is unclear whether, or how, trade publishers will use personal data to predict and assess consumers' habits, but we know that ebook platforms are using personal data to assess customers' reading habits and for "more like this" advertising and other marketing and tools used to track and assess consumer preferences and trends.

In light of the platforms' data collections and use, and the potential for publishers to similarly use purchasers' data, lawmakers should ensure that data privacy laws protect consumers that use ebook platforms.

Library professionals and legal scholars have carved out a special privacy right for people as they read and research.<sup>129</sup> Librarian ethics include "intellectual freedom," the principle of a private right to read and research based on the idea that people should be able to read, learn, and debate without being monitored and recorded.<sup>130</sup> The right of every individual to both seek and receive information from all points of view without restriction is inherent in First Amendment freedoms of speech, association, and assembly. When people know that their research and reading habits are being tracked and recorded, it has a chilling effect on people's ability to seek information. Most states have laws that protect people's library records as confidential,<sup>131</sup> and those laws should be updated to ensure that they cover ebook platforms. These laws codify the American Library Association's best practices to protect intellectual freedom, including keeping circulation records and other patron-identifying records confidential.<sup>132</sup> Librarians also lobbied state governments to pass library patron privacy laws to safeguard print book readers' privacy, convincing every state to pass a library-patron privacy law. (The two states that lack privacy laws have attorney generals' opinions guaranteeing library privacy.)<sup>133</sup>

---

129 Neil Richards, *Intellectual Privacy* (2015); *Intellectual Freedom Manual* (Martin Garnar, Trina Magi, and American Libr. Ass'n Off. for Intellectual Freedom, eds., 10th ed. 2021).

130 The American Library Association's (ALA) intellectual freedom manual describes an information seeker's privacy right as "the right to open inquiry without having the subject of one's interest examined or scrutinized by others." *Id.* at 177.

131 Forty-eight states have library privacy statutes, and the two states that do not have library privacy laws have attorney general opinions that require libraries to maintain their patrons' privacy. *State Privacy Laws Regarding Library Records*, ALA, <https://www.ala.org/advocacy/privacy/statelaws> (Nov. 2021). For discussion of the history of these laws, see Rhoda Garoogian, Librarian/Patron Confidentiality: An Ethical Challenge, 40 *Library Trends* 216–33 (1991); Bruce M. Kennedy, Confidentiality of Library Records: A Survey of Problems, Policies, and Laws, 81 *Law Libr. J.* 733–67.

132 Library Bill of Rights VII, <https://www.ala.org/advocacy/intfreedom/librarybill>

133 Kennedy, *supra* note 103.

These library laws were passed long before ebooks and the ebook platform industry existed. The lawmakers and librarians that drafted library privacy legislation did not imagine our current ebook systems, and they did not predict that personal data analytics would become a prominent feature of book marketing or a copyright-protection measure on most digital information platforms. Many states have updated their library privacy laws to include email records and electronic book circulation trackers (associated with library card barcodes, etc.), but few have been amended to account for internet-connected ebook platforms and ereaders. Librarians have raised their concerns about the threats ebook platforms pose to their patrons' intellectual freedom, but state legislators have yet to update the laws to reflect those concerns.<sup>134</sup>


As we wait for data privacy legislation in Congress, states could update their library privacy laws to increase their scope, if they haven't already done so. Today's library privacy laws should protect as confidential records all of the personal data collected in libraries, including data collected by third-party vendors, as they do paper and other electronic library records.



---

134 When the Amazon Kindle ebooks platform was introduced in libraries, the American Library Association's director of the Office of Intellectual Freedom responded with concerns about library patrons' privacy, saying, "The right to read freely depends upon the knowledge that what one is reading is not monitored or tracked." Deborah Caldwell-Stone, A Digital Dilemma: Ebooks and Users' Rights, *American Libraries* (May 29, 2012), <https://americanlibrariesmagazine.org/2012/05/29/a-digital-dilemma-ebooks-and-users-rights/>

 **Engelberg Center on Innovation Law & Policy**  
139 MacDougal Street Room 408 | New York, NY 10012

 [NYUEngelberg.org](https://nyuengelberg.org)

 [engelberg.center@nyu.edu](mailto:engelberg.center@nyu.edu)

 [@NYUEngelberg](https://twitter.com/NYUEngelberg)

