

## FAIR USE WEEK 2023: LOOKING BACK AT GOOGLE BOOKS EIGHT YEARS LATER

Posted February 24, 2023

*This post is authored by Authors Alliance Senior Staff Attorney, Rachel Brooke.*

More recent members and readers may not be aware that Authors Alliance was founded in the wake of *Authors Guild v. Google*, a class action fair use case in the Second Circuit that was litigated for nearly a decade, and finally resolved in favor of Google in 2015. The case concerned the Google Books project—an initiative launched by Google whereby the company partnered with university libraries to scan books in their collections. These scans would ultimately be made available as a full-text searchable database for the public to search through for particular terms, with short “snippets” displayed accompanying the search results. Users could not, however, view or read the scanned books in their entirety. The Authors Guild, along with several authors, filed a lawsuit against Google alleging that scanning the books and displaying these snippets constituted copyright infringement.

In addition to Authors Guild representing its members in the litigation, its associated plaintiffs brought the case as a class action, claiming to bring the case on behalf of a broad group of authors: “[a]ll persons residing in the United States who hold a United States copyright interest in one or more Books reproduced by Google as part of its Library Project” who were either authors or the authors’ heirs.

But many of these authors did *not* agree with the Authors Guild’s stance in the case, and felt that the Google Books project served their interests in sharing knowledge, seeing their creations be preserved, and reaching readers interested in their work. A group of authors and scholars came together to share their views with the district court, many of whom would soon become founding members of Authors Alliance. Many of those same authors signed on to amicus briefs before both the district court and Second Circuit explaining why they opposed the litigation and supported Google’s fair use defense. Then, in 2014, Authors Alliance submitted its first amicus brief to the Second Circuit, supporting Google’s ultimately successful fair use defense. The plaintiffs later appealed the Second Circuit’s ruling, asking the Supreme Court to weigh in, but the Court ultimately declined to hear the case, leaving the Second Circuit’s ruling intact.

Nearly a decade later, the effects of *Google Books* can still be seen in fair use decisions and copyright policy developments involving the challenges of adapting copyright to the digital world. In today's post, I'll reflect on how *Google Books* can be contextualized within today's fair use landscape and share my thoughts on what the case can tell us about copyright in the digital world.

## *Google Books* and Transformativeness

A major question in *Authors Guild v. Google* was whether Google's use of the copyrighted works was "transformative," a key component of the fair use inquiry. When a use is found to be transformative, this in practice weighs heavily in favor of a finding of fair use. In the case, the court found that Google's scanning, as well as the search and snippet display functions, were transformative because the service "augments public knowledge by making available information *about* [the] books without providing the public with a substantial substitute for . . . the original works." This was because Google Books provided information about the books—such as the author and publisher information—without creating substitutes of the original works. In other words, readers could learn about the books they searched through, but could not read the books in full—to do this, those readers would have to purchase or borrow copies through the normal channels.

Since the doctrine of transformativeness was established in the 1994 landmark Supreme Court case, *Campbell v. Acuff-Rose Music*, there have been myriad questions about the precise contours of what it means for a use to be transformative. *Campbell* established that a use is transformative when it endows the secondary work with a "new meaning or message," but it can be difficult to apply this test in practice, particularly in the context of new or nascent technologies. *Google Books* tells us that scanning works in order to create a full-text searchable database with limited snippet displays is a transformative use based on its new and different purpose from the purpose of the works themselves. Furthermore, it reinforces the notion that a use is particularly likely to be considered transformative when it serves the underlying purpose of copyright law: incentivizing new creation for the benefit of the public and "enriching public knowledge." By highlighting that Google contributed to public knowledge about books through its scanning activities and the Google Books search function, the court helped bring fair use for scholarship and research—two key prototypical uses

established in the 1976 Copyright Act—into the digital age, setting an important precedent for later cases.

## *Google Books* and Derivative Works

One of the plaintiffs' arguments in *Google Books* was that Google's full-text searchable database constituted a derivative work. One of a copyright holder's exclusive rights is the right to prepare derivative works—such as adaptations, abridgements, or translations of the original work—and the plaintiffs alleged that this right had been infringed. The court disagreed, finding that Google's use had a transformative purpose, whereas derivative works tend to involve a transformation in form, such as the adaptation of a novel into a movie or an audiobook. Furthermore, the court explained that derivative works are "those that re-present the protected aspects of the original work, i.e., its expressive content, converted into an altered form[.]" In contrast, the Google Books project provided information about the books and offered a limited "snippet" view, but did not re-present the expressive content: the full text of the books themselves.

The distinction the court drew between transformative fair uses and derivative works in *Google Books* is an important one, as it can often be a close question whether a work involves a transformative purpose or merely represents the same work in a new form, without enough added to tip the scales towards fair use. And it is a question that continues to arise in fair use cases today: just last year, the Supreme Court agreed to hear *Warhol Foundation v. Goldsmith*, a case about whether Andy Warhol's creation of a series of screenprints of the late musical artist Prince which drew from a photograph taken by photographer Lynn Goldsmith qualified as a fair use. We've covered this case extensively on our blog over the past few years, and submitted an amicus brief in the case. Our brief argues (among other things) that Warhol's screen prints involve much more than a transformation in form: they are stylistically and visually distinct from Goldsmith's photograph, and endow the photograph with a new meaning or message, making the use highly transformative.

As in *Google Books*, the parties and amici in *Goldsmith* grapple with the line between transformative uses and the creation of derivative works, an often complicated and fact-sensitive determination. In this context, *Google Books* serves as a reminder that

fair use is not a one-size-fits-all determination. Yet it also provides support for arguments advanced by Authors Alliance and others that simply because a transformation in form exists—in the *Google Books* case, the transformation from a print book to a scanned copy, and in *Goldsmith*, the transformation of a black and white photo to a series of colorful screenprints—does not mean that a secondary use cannot be a fair one. Warhol’s use did not merely “re-present the protected aspects of the original work[’s] . . . expressive content,” but was transformative in the different “purpose, character, expression, meaning, and message” it conveyed.

### *Google Books* and Controlled Digital Lending

The practice of **controlled digital lending** (“CDL”)—and the arguments in favor of it constituting a fair use—can be traced back in part to the fair use principles established and reinforced in *Google Books*. As I argue in **our amicus brief** in *Hachette Books v. Internet Archive*, a case about—among other things—whether CDL constitutes a fair use, *Google Books* shows that copying the entirety of a work in the process of making a transformative use of it can be fully consistent with fair use.

Another important suggestion in the *Google Books* case, made at the district court level, was that the Google Books search function could actually drive book sales: the search results were accompanied by links to purchase the book, and research suggested that this could enhance sales of those books. This is analogous to the effects of library lending: library readers **often purchase books by authors they first discovered at the library**, an effect which can apply with equal force when the library patron borrows a CDL scan. Indeed, several other **amici** in *Hachette Books* argue that the finding that the Google Books search was a fair use lent substantial support for the argument that CDL is a fair use, based on both the factual similarities between the two initiatives and their shared objective of “enriching public knowledge.”

As in *Google Books*, CDL also helps authors reach readers who could not otherwise access their books, and achieves this through scanning books on library shelves. And also like *Google Books*, CDL helps solve the problem of **20th century works “disappearing”**: the commercial life of a book tends to be much shorter than the term of copyright, so when books under copyright go out of print, they can

disappear into obscurity. Scanning these books to preserve them ensures that the knowledge they advance will not be lost.

## *Google Books* and Text Data Mining

Text data mining—the process of using automated techniques aimed at quantitatively analyzing text and other data—is also widely considered to be a fair use, and this determination is similarly built in part on the building blocks established in *Google Books*. As was the case in *Google Books*, the results of text data mining research provide information *about* the works being studied, and cannot in any way serve as substitutes for the content of the works. In fact, one important aspect of the new exemption to DMCA liability for text data mining, which **Authors Alliance successfully petitioned for in 2021**, is that researchers are not able to use the works in the text data mining corpus for consumptive purposes. And also like *Google Books*, researchers *are* able to view the content in a limited manner to verify their findings, analogous to Google Books’s snippet view. The new TDM exemption was a huge win for Authors Alliance members, and something to celebrate for all scholars engaged in this important research. Importantly, the precedent established by Google Books strongly supported its adoption and the Register of Copyright’s suggestion that text data mining was **likely to be a fair use**.

## Looking Forward: *Google Books* and Artificial Intelligence

In recent years, scholars and researchers have grappled with the implications of copyright protection on AI-generated content and AI models more generally. **The holding in *Google Books* provides some support for companies’ and researchers’ ability to engage in these activities: one important factor in the case was that Google Books did not harm the market for the books at issue in the case, since the books in the database could not serve as substitutes for the books themselves. Similarly, when copyrighted works are used to train AI, the output cannot serve as a substitute for the copyrighted works, and the market for those works is not harmed, even if—like the plaintiffs in *Google Books*—the copyright holders might prefer that their works not be used in this way. *Google Books* establishes that simply because copyrighted works are used as “input” in a given model, this does not mean that the *outputs* constitute infringement.** It is also worth noting that the court found Google’s use to be fair despite the fact that it was a use by a commercial,

profit-seeking entity. While a commercial use can sometimes tip the scales in favor of finding a use to not be fair, this can be overcome by a socially beneficial, transformative purpose. This could arguably apply with equal force to AI models trained on copyrighted works which contribute to our understanding of the world, despite the fact that commercial entities are often the ones deploying these technologies.

Eight years after it was decided, the legacy of *Google Books* endures in policy debates and copyright lawsuits that capture the public's attention. Policymakers and judges would be wise to heed the lessons it teaches about the value of advancing public knowledge through digitization and the use of copyrighted works for new and socially beneficial purposes. As we await policy developments regarding text data mining and wait for decisions in *Goldsmith* and *Hachette Books*, it is my hope that this legacy will live on, reminding us all of the vast capabilities of information technology to enrich our understanding of the world and advance the progress of knowledge, which, after all, is what copyright law is all about.

Share this post [Share on Twitter](#)[Share on Facebook](#)[Share on LinkedIn](#)