

# THE AUTHORS GUILD, INC. v. GOOGLE INC.

United States District Court, S.D. New York.

954 F.Supp.2d 282

November 14, 2013.

## 12 DENNY CHIN, Circuit District Judge.

13 Since 2004, when it announced agreements with several major research libraries to digitally copy books in their collections, defendant Google Inc. ("Google") has scanned more than twenty million books. It has delivered digital copies to participating libraries, created an electronic database of books, and made text available for online searching through the use of "snippets." Many of the books scanned by Google, however, were under copyright, and Google did not obtain permission from the copyright holders for these usages of their copyrighted works. As a consequence, in 2005, plaintiffs brought this class action charging Google with copyright infringement.

14 Before the Court are the parties' cross-motions for summary judgment with respect to Google's defense of fair use under § 107 of the Copyright Act, 17 U.S.C. § 107. For the reasons set forth below, Google's motion for summary judgment is granted and plaintiffs' motion for partial summary judgment is denied. Accordingly, judgment will be entered in favor of Google dismissing the case.

## BACKGROUND

### A. The Facts

17 For purposes of this motion, the facts are not in dispute. [...] They are summarized as follows:

#### 18 1. The Parties

19 Plaintiff Jim Bouton, the former pitcher for the New York Yankees, is the legal or beneficial owner of the U.S. copyright in the book *Ball Four*. Plaintiff Betty Miles is the legal or beneficial owner of the U.S. copyright in the book *The Trouble with Thirteen*. Plaintiff Joseph Goulden is the legal or beneficial owner of the U.S. copyright in the book *The Superlawyers: The Small and Powerful World of the Great Washington Law Firms*. [...] [2] All three books have been scanned by Google and are available for search on Google's website, without plaintiffs' permission. [...] Plaintiff The Authors Guild, Inc., is the nation's largest organization of published authors and it advocates for and supports the copyright and contractual interests of published writers. [...]

20 Google owns and operates the largest Internet search engine in the world. [...] Each day, millions of people use Google's search engine free of charge; commercial and other entities pay to display ads on Google's websites and on other websites that contain Google ads. [...] Google is a for-profit entity, and for the year ended December 31, 2011, it reported over \$36.5 billion in advertising revenues. [...]

21 **2. The Google Books Project**

- 22 In 2004, Google announced two digital books programs. The first, initially called "Google Print" and later renamed the "Partner Program," involved the "hosting" and display of material provided by book publishers or other rights holders. [...] The second became known as the "Library Project," and over time it involved the digital scanning of books in the collections of the New York Public Library, the Library of Congress, and a number of university libraries. [...]
- 23 The Partner Program and the Library Project together comprise the Google Books program ("Google Books"). [...] All types of books are encompassed, including novels, biographies, children's books, reference works, textbooks, instruction manuals, treatises, dictionaries, cookbooks, poetry books, and memoirs. [...] Some 93% of the books are non-fiction while approximately 7% are fiction.[...] Both in-print and out-of-print books are included, although the great majority are out-of-print. [...]
- 24 In the Partner Program, works are displayed with permission of the rights holders. [...] The Partner Program is aimed at helping publishers sell books and helping books become discovered. [...] Initially, Google shared revenues from ads with publishers or other rights holders in certain circumstances. In 2011, however, Google stopped displaying ads in connection with all books. [...] Partners provide Google with a printed copy of their books for scanning, or a digital copy if one already exists. [...] Partners decide how much of their books — from a few sample pages to the entire book — are browsable. [...] As of early 2012, the Partner Program included approximately 2.5 million books, with the consent of some 45,000 rights holders. [...]
- 25 As for the Library Project, Google has scanned more than twenty million books, in their entirety, using newly-developed scanning technology. [...] Pursuant to their agreement with Google, participating libraries can download a digital copy of each book scanned from their collections. [...] Google has provided digital copies of millions of these books to the libraries, in accordance with these agreements. [...] Some libraries agreed to allow Google to scan only public domain works, while others allowed Google to scan in-copyright works as well. [...]
- 26 Google creates more than one copy of each book it scans from the library collections, and it maintains digital copies of each book on its servers and back-up tapes. [...] Participating libraries have downloaded digital copies of in-copyright books scanned from their collections. [...] They may not obtain a digital copy created from another library's book. [...] The libraries agree to abide by the copyright laws with respect to the copies they make. [...]
- 27 Google did not seek or obtain permission from the copyright holders to digitally copy or display verbatim expressions from in-copyright books. [...] Google has not compensated copyright holders for its copying of or displaying of verbatim expression from in-copyright books or its making available to libraries for downloading of digital copies of in-copyright books scanned from their collections. [...]

28 **3. Google Books**

29 In scanning books for its Library Project, including in-copyright books, Google uses optical character recognition technology to generate machine-readable text, compiling a digital copy of each book. [...] Google analyzes each scan and creates an overall index of all scanned books. The index links each word or phrase appearing in each book with all of the locations in all of the books in which that word or phrase is found. The index allows a search for a particular word or phrase to return a result that includes the most relevant books in which the word or phrase is found. [...] Because the full texts of books are digitized, a user can search the full text of all the books in the Google Books corpus. [...]

30 Users of Google's search engine may conduct searches, using queries of their own design. [...] In response to inquiries, Google returns a list of books in which the search term appears. [...] A user can click on a particular result to be directed to an "About the Book" page, which will provide the user with information about the book in question. The page includes links to sellers of the books and/or libraries that list the book as part of their collections. No advertisements have ever appeared on any About the Book page that is part of the Library Project. [...]

31 For books in "snippet view" (in contrast to "full view" books), Google divides each page into eighths — each of which is a "snippet," a verbatim excerpt. [...] Each search generates three snippets, but by performing multiple searches using different search terms, a single user may view far more than three snippets, as different searches can return different snippets. [...] For example, by making a series of consecutive, slightly different searches of the book *Ball Four*, a single user can view many different snippets from the book. [...]

32 Google takes security measures to prevent users from viewing a complete copy of a snippet-view book. For example, a user cannot cause the system to return different sets of snippets for the same search query; the position of each snippet is fixed within the page and does not "slide" around the search term; only the first responsive snippet available on any given page will be returned in response to a query; one of the snippets on each page is "black-listed," meaning it will not be shown; and at least one out of ten entire pages in each book is black-listed. [...] An "attacker" who tries to obtain an entire book by using a physical copy of the book to string together words appearing in successive passages would be able to obtain at best a patchwork of snippets that would be missing at least one snippet from every page and 10% of all pages. [...] In addition, works with text organized in short "chunks," such as dictionaries, cookbooks, and books of haiku, are excluded from snippet view. [...]

33 **4. The Benefits of the Library Project and Google Books**

34 The benefits of the Library Project are many. First, Google Books provides a new and efficient way for readers and researchers to find books. [...] It makes tens of millions of books searchable by words and phrases. It provides a searchable index linking each word in any book to all books in which that word appears. [...] Google Books has become an essential research tool, as it helps librarians identify and find research sources, it makes

the process of interlibrary lending more efficient, and it facilitates finding and checking citations. [...] Indeed, Google Books has become such an important tool for researchers and librarians that it has been integrated into the educational system — it is taught as part of the information literacy curriculum to students at all levels. [...]

- 35 Second, in addition to being an important reference tool, Google Books greatly promotes a type of research referred to as "data mining" or "text mining." [...] Google Books permits humanities scholars to analyze massive amounts of data — the literary record created by a collection of tens of millions of books. Researchers can examine word frequencies, syntactic patterns, and thematic markers to consider how literary style has changed over time. [...] Using Google Books, for example, researchers can track the frequency of references to the United States as a single entity ("the United States is") versus references to the United States in the plural ("the United States are") and how that usage has changed over time. [...] The ability to determine how often different words or phrases appear in books at different times "can provide insights about fields as diverse as lexicography, the evolution of grammar, collective memory, the adoption of technology, the pursuit of fame, censorship, and historical epidemiology." [...] Third, Google Books expands access to books. In particular, traditionally underserved populations will benefit as they gain knowledge of and access to far more books. Google Books provides print-disabled individuals with the potential to search for books and read them in a format that is compatible with text enlargement software, text-to-speech screen access software, and Braille devices. Digitization facilitates the conversion of books to audio and tactile formats, increasing access for individuals with disabilities. [...] Google Books facilitates the identification and access of materials for remote and underfunded libraries that need to make efficient decisions as to which resources to procure for their own collections or through interlibrary loans. [...]
- 36 Fourth, Google Books helps to preserve books and give them new life. Older books, many of which are out-of-print books that are falling apart buried in library stacks, are being scanned and saved. [...] These books will now be available, at least for search, and potential readers will be alerted to their existence.
- 37 Finally, by helping readers and researchers identify books, Google Books benefits authors and publishers. When a user clicks on a search result and is directed to an "About the Book" page, the page will offer links to sellers of the book and/or libraries listing the book as part of their collections. [...] The About the Book page for *Ball Four*, for example, provides links to Amazon.com, Barnes&Noble.com, Books-A-Million, and IndieBound. [...] A user could simply click on any of these links to be directed to a website where she could purchase the book. Hence, Google Books will generate new audiences and create new sources of income.
- 38 As amici observe: "Thanks to . . . [Google Books], librarians can identify and efficiently sift through possible research sources, amateur historians have access to a wealth of previously obscure material, and everyday readers and researchers can find books that were once buried in research library archives." [...]

## B. Procedural History

- 40 Plaintiffs commenced this action on September 20, 2005, alleging, inter alia, that Google committed copyright infringement by scanning copyrighted books and making them available for search without permission of the copyright holders. From the outset, Google's principal defense was fair use under § 107 of the Copyright Act[...].
- 41 After extensive negotiations, the parties entered into a proposed settlement resolving plaintiffs' claims on a class-wide basis. On March 22, 2011, I issued an opinion rejecting the proposed settlement on the grounds that it was not fair, adequate, and reasonable. [...]
- 42 Thereafter, the parties engaged in further settlement discussions, but they were unable to reach agreement. The parties proposed and I accepted a schedule that called for the filing of plaintiffs' class certification motion, the completion of discovery, and then the filing of summary judgment motions. [...] Plaintiffs filed a fourth amended class action complaint (the "Complaint") on October 14, 2011. [...] [...]
- 43 Plaintiffs filed their class certification motion and Google filed its motion to dismiss the Authors Guild's claims. On May 31, 2012, I issued an opinion denying Google's motion to dismiss and granting the individual plaintiffs' motion for class certification. [...]
- 44 [...] The parties thereafter filed the instant cross-motions for summary judgment. Before the motions were fully submitted, however, the Second Circuit issued an order on September 17, 2012, staying these proceedings pending an interlocutory appeal by Google from my decision granting class certification. [...]
- 45 On July 1, 2013, without deciding the merits of the appeal, the Second Circuit vacated my class certification decision, concluding that "resolution of Google's fair use defense in the first instance will necessarily inform and perhaps moot our analysis of many class certification issues." [...] The Second Circuit remanded the case "for consideration of the fair use issues." [...]

## DISCUSSION

- 48 For purposes of these motions, I assume that plaintiffs have established a prima facie case of copyright infringement against Google under 17 U.S.C. § 106. [...] Google has digitally reproduced millions of copyrighted books, including the individual plaintiffs' books, maintaining copies for itself on its servers and backup tapes. See 17 U.S.C. § 106(1) (prohibiting unauthorized reproduction). Google has made digital copies available for its Library Project partners to download. See 17 U.S.C. § 106(3) (prohibiting unauthorized distribution). Google has displayed snippets from the books to the public. See 17 U.S.C. § 106(5) (prohibiting unauthorized display). Google has done all of this, with respect to in-copyright books in the Library Project, without license or permission from the copyright owners. The sole issue now before the Court is whether Google's use of the copyrighted works is "fair use" under the copyright laws. For the reasons set forth below, I conclude that it is.

## A. Applicable Law

- 50 Fair use is a defense to a claim of copyright infringement. The doctrine permits the fair use of copyrighted works "to fulfill copyright's very purpose, t[t]o promote the Progress of Science and useful Arts." *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 575 (1994) (quoting U.S. Const., Art. I, § 8, cl. 8))[...]. Copyright law seeks to achieve that purpose by providing sufficient protection to authors and inventors to stimulate creative activity, while at the same time permitting others to utilize protected works to advance the progress of the arts and sciences. [...] As the Supreme Court has held, "[f]rom the infancy of copyright protection, some opportunity for fair use of copyrighted materials has been thought necessary to fulfill copyright's very purpose." [...]
- 51 The fair use doctrine is codified in § 107 of the Copyright Act, which provides in relevant part as follows:
- 52 [T]he fair use of a copyrighted work, . . . for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include —
- 53 (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- 54 (2) the nature of the copyrighted work;
- 55 (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- 56 (4) the effect of the use upon the potential market for or value of the copyrighted work.[...]
- 58 The determination of fair use is "an open-ended and context-sensitive inquiry," [...] and thus the fair use doctrine calls for "case-by-case analysis," [...]. The four factors enumerated in the statute are non-exclusive and provide only "general guidance"; they are to be explored and weighed together, "in light of the purposes of copyright." [...] As fair use is an affirmative defense to a claim of copyright infringement, the proponent carries the burden of proof as to all issues in dispute. [...]
- 59 A key consideration is whether, as part of the inquiry into the first factor, the use of the copyrighted work is "transformative," that is, whether the new work merely "supersedes" or "supplants" the original creation, or whether it:
- 60 instead adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message; it asks, in other words, whether and to what extent the new work is "transformative."
- 61 *Campbell*, 510 U.S. at 579 (quoting *Leval, Toward a Fair Use Standard*, 103 *Harv. L. Rev.* at 1111)[...] Although transformative use is not "absolutely necessary" to a finding of fair



use, "the goal of copyright, to promote science and the arts, is generally furthered by the creation of transformative works." [...]

## B. Application

63 I discuss each of the four factors separately, and I then weigh them together.

### 64 1. Purpose and Character of Use

65 The first factor is "the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes." 17 U.S.C. § 107(1).

66 Google's use of the copyrighted works is highly transformative. Google Books digitizes books and transforms expressive text into a comprehensive word index that helps readers, scholars, researchers, and others find books. Google Books has become an important tool for libraries and librarians and cite-checkers as it helps to identify and find books. The use of book text to facilitate search through the display of snippets is transformative. See *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1168 (9th Cir. 2007) (holding that use of works — "thumbnail images," including copyrighted photographs — to facilitate search was "transformative"); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003) (same); see also *Bill Graham Archives*, 448 F.3d at 609-11 (holding that display of images of posters in 480-page cultural history of the Grateful Dead was transformative, explaining that "[w]hile the small size [of the images of the posters] is sufficient to permit readers to recognize the historical significance of the posters, it is inadequate to offer more than a glimpse of their expressive value"). The display of snippets of text for search is similar to the display of thumbnail images of photographs for search or small images of concert posters for reference to past events, as the snippets help users locate books and determine whether they may be of interest. Google Books thus uses words for a different purpose — it uses snippets of text to act as pointers directing users to a broad selection of books.

67 Similarly, Google Books is also transformative in the sense that it has transformed book text into data for purposes of substantive research, including data mining and text mining in new areas, thereby opening up new fields of research. Words in books are being used in a way they have not been used before. Google Books has created something new in the use of book text — the frequency of words and trends in their usage provide substantive information.

68 Google Books does not supersede or supplant books because it is not a tool to be used to read books. Instead, it "adds value to the original" and allows for "the creation of new information, new aesthetics, new insights and understandings." [...] Hence, the use is transformative.

69 It is true, of course, as plaintiffs argue, that Google is a for-profit entity and Google Books is largely a commercial enterprise. The fact that a use is commercial "tends to weigh against a finding of fair use." [...] On the other hand, fair use has been found even where a defendant benefitted commercially from the unlicensed use of copyrighted works. [...] Here, Google does not sell the scans it has made of books for Google Books; it does

not sell the snippets that it displays; and it does not run ads on the About the Book pages that contain snippets. It does not engage in the direct commercialization of copyrighted works. [...] Google does, of course, benefit commercially in the sense that users are drawn to the Google websites by the ability to search Google Books. While this is a consideration to be acknowledged in weighing all the factors, even assuming Google's principal motivation is profit, the fact is that Google Books serves several important educational purposes.

70 Accordingly, I conclude that the first factor strongly favors a finding of fair use.

## 71 **2. Nature of Copyrighted Works**

72 The second factor is "the nature of the copyrighted work." 17 U.S.C. § 107(2).[...] Here, the works are books — all types of published books, fiction and non-fiction, in-print and out-of-print. While works of fiction are entitled to greater copyright protection, [...] here the vast majority of the books in Google Books are non-fiction. Further, the books at issue are published and available to the public. These considerations favor a finding of fair use. [...]

## 73 **3. Amount and Substantiality of Portion Used**

74 The third factor is "the amount and substantiality of the portion used in relation to the copyrighted work as a whole." 17 U.S.C. § 107(3). Google scans the full text of books — the entire books — and it copies verbatim expression. On the other hand, courts have held that copying the entirety of a work may still be fair use. [...] Here, as one of the keys to Google Books is its offering of full-text search of books, full-work reproduction is critical to the functioning of Google Books. Significantly, Google limits the amount of text it displays in response to a search.

75 On balance, I conclude that the third factor weighs slightly against a finding of fair use.

## 76 **4. Effect of Use Upon Potential Market or Value**

77 The fourth factor is "the effect of the use upon the potential market for or value of the copyrighted work." 17 U.S.C. § 107(4). Here, plaintiffs argue that Google Books will negatively impact the market for books and that Google's scans will serve as a "market replacement" for books. [...] It also argues that users could put in multiple searches, varying slightly the search terms, to access an entire book. [...]

78 Neither suggestion makes sense. Google does not sell its scans, and the scans do not replace the books. While partner libraries have the ability to download a scan of a book from their collections, they owned the books already — they provided the original book to Google to scan. Nor is it likely that someone would take the time and energy to input countless searches to try and get enough snippets to comprise an entire book. Not only is that not possible as certain pages and snippets are blacklisted, the individual would have to have a copy of the book in his possession already to be able to piece the different snippets together in coherent fashion.

79 To the contrary, a reasonable factfinder could only find that Google Books enhances the



sales of books to the benefit of copyright holders. An important factor in the success of an individual title is whether it is discovered — whether potential readers learn of its existence. [...] Google Books provides a way for authors' works to become noticed, much like traditional in-store book displays. [...] Indeed, both librarians and their patrons use Google Books to identify books to purchase. [...] Many authors have noted that online browsing in general and Google Books in particular helps readers find their work, thus increasing their audiences. Further, Google provides convenient links to booksellers to make it easy for a reader to order a book. In this day and age of on-line shopping, there can be no doubt but that Google Books improves books sales.

80 Hence, I conclude that the fourth factor weighs strongly in favor of a finding of fair use.

## 81 **5. Overall Assessment**

82 Finally, the various non-exclusive statutory factors are to be weighed together, along with any other relevant considerations, in light of the purposes of the copyright laws.

83 In my view, Google Books provides significant public benefits. It advances the progress of the arts and sciences, while maintaining respectful consideration for the rights of authors and other creative individuals, and without adversely impacting the rights of copyright holders. It has become an invaluable research tool that permits students, teachers, librarians, and others to more efficiently identify and locate books. It has given scholars the ability, for the first time, to conduct full-text searches of tens of millions of books. It preserves books, in particular out-of-print and old books that have been forgotten in the bowels of libraries, and it gives them new life. It facilitates access to books for print-disabled and remote or underserved populations. It generates new audiences and creates new sources of income for authors and publishers. Indeed, all society benefits.

84 Similarly, Google is entitled to summary judgment with respect to plaintiffs' claims based on the copies of scanned books made available to libraries. Even assuming plaintiffs have demonstrated a prima facie case of copyright infringement, Google's actions constitute fair use here as well. Google provides the libraries with the technological means to make digital copies of books that they already own. The purpose of the library copies is to advance the libraries' lawful uses of the digitized books consistent with the copyright law. The libraries then use these digital copies in transformative ways. They create their own full-text searchable indices of books, maintain copies for purposes of preservation, and make copies available to print-disabled individuals, expanding access for them in unprecedented ways. Google's actions in providing the libraries with the ability to engage in activities that advance the arts and sciences constitute fair use.

85 To the extent plaintiffs are asserting a theory of secondary liability against Google, the theory fails because the libraries' actions are protected by the fair use doctrine. Indeed, in the HathiTrust case, Judge Baer held that the libraries' conduct was fair use. See *Authors Guild, Inc. v. HathiTrust*, 902 F. Supp. 2d 445, 460-61, 464 (S.D.N.Y. 2012) ("I cannot imagine a definition of fair use that would not encompass the transformative uses made by Defendants' [Mass Digitization Project] and would require that I terminate this invaluable contribution to the progress of science and cultivation of the arts that at the

same time effectuates the ideals espoused by the [Americans with Disabilities Act]."). The fair use analysis set forth above with respect to Google Books applies here as well to the libraries' use of their scans, and if there is no liability for copyright infringement on the libraries' part, there can be no liability on Google's part.

### **CONCLUSION**

87 For the reasons set forth above, plaintiffs' motion for partial summary judgment is denied and Google's motion for summary judgment is granted. Judgment will be entered in favor of Google dismissing the Complaint. Google shall submit a proposed judgment, on notice, within five business days hereof.[...]