13-4829-cv
Authors Guild v. Google, Inc.

UNITED STATES COURT OF APPEALS
FOR THE SECOND CIRCUIT

August Term, 2014

(Argued: December 3, 2014 Decided: October 16, 2015)

Docket No. 13-4829-cv

- -X

THE AUTHORS GUILD, BETTY MILES, JIM BOUTON, JOSEPH GOULDEN, individually and on behalf of all others similarly situated,
Plaintiff-Appellants,

HERBERT MITGANG, DANIEL HOFFMAN, individually and on behalf of all others similarly situated, PAUL DICKSON, THE MCGRAW-HILL COMPANIES, INC., PEARSON EDUCATION, INC., SIMON & SCHUSTER, INC., ASSOCIATION OF AMERICAN PUBLISHERS, INC., CANADIAN STANDARD ASSOCIATION, JOHN WILEY & SONS, INC., individually and on behalf of all others similarly situated,

Plaintiffs,

v.

GOOGLE, INC.,
Defendant-Appellee.

- X

Before: LEVAL, CABRANES, PARKER, *Circuit Judges*.

13-4829-cv
Authors Guild v. Google, Inc.

Plaintiff-appellants, who are authors of published books under copyright, appeal from the judgment of the United States District Court for the Southern District of New York (Chin, *J.*) in favor of Defendant Google, Inc. Plaintiffs sued Google, alleging that its Library Project and Google Books project infringe Plaintiffs' copyrights. Through these projects, Google makes and retains digital copies of books submitted to it by major libraries, allows the libraries that submitted a book to download and retain a digital copy, and allows the public to search the texts of the digitally copied books and see displays of snippets of text. The district court granted summary judgment based on its conclusion that Google's copying is fair use under 17 U.S.C. § 107 and is therefore not infringing. The Court of Appeals concludes that the defendant's copying is transformative within the meaning of *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578-585 (1994), does not offer the public a meaningful substitute for matter protected by the plaintiffs' copyrights, and satisfies § 107's test for fair use. AFFIRMED.

For Plaintiff-Appellants: PAUL M. SMITH, JENNER & BLOCK LLP, WASHINGTON, DC (Edward H. Rosenthal, Jeremy S. Goldman, Anna Kadyshevich, Andrew D. Jacobs, Frankfurt Kurnit Klein & Selz PC, New York, NY *on the brief*)

For Defendant-Appellee: SETH P. WAXMAN, WILMER CUTLER PICKERING HALE AND DORR LLP, WASHINGTON, D.C. (Louis R. Cohen, Daniel P. Kearney, Jr., Weili J. Shaw, Wilmer Cutler Pickering Hale and Dorr LLP, Washington D.C & Daralyn J. Durie, Joseph C. Gratz, Durie Tangri LLP, San Francisco, CA *on the brief*)

Leval, *Circuit Judge*:

This copyright dispute tests the boundaries of fair use. Plaintiffs, who are authors of published books under copyright, sued Google, Inc. ("Google") for copyright infringement in the United States District Court for the Southern District of New York (Chin, *J.*). They appeal from the grant of summary judgment in Google's favor. Through its Library Project and its Google Books project, acting without permission of rights holders, Google has made digital copies of tens of millions of books, including Plaintiffs', that were submitted to it for that purpose by major libraries. Google has scanned the digital copies and established a publicly available search function. An Internet user can use this function to search without charge to determine whether

the book contains a specified word or term and also see "snippets" of text containing the searched-for terms. In addition, Google has allowed the participating libraries to download and retain digital copies of the books they submit, under agreements which commit the libraries not to use their digital copies in violation of the copyright laws. These activities of Google are alleged to constitute infringement of Plaintiffs' copyrights. Plaintiffs sought injunctive and declaratory relief as well as damages.

Google defended on the ground that its actions constitute "fair use," which, under 17 U.S.C. § 107, is "not an infringement." The district court agreed. *Authors Guild, Inc. v. Google Inc.*, 954 F. Supp. 2d 282, 294 (S.D.N.Y. 2013). Plaintiffs brought this appeal.

Plaintiffs contend the district court's ruling was flawed in several respects. They argue: (1) Google's digital copying of entire books, allowing users through the snippet function to read portions, is not a "transformative use" within the meaning of *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578-585 (1994), and provides a substitute for Plaintiffs' works; (2) notwithstanding that Google provides public access to the search and snippet functions without charge and without advertising, its ultimate commercial profit motivation and its derivation of revenue from its dominance of the world-wide Internet search market to which the books project contributes, preclude a finding of fair use; (3) even if Google's copying and revelations of text do not infringe plaintiffs' *books*, they infringe Plaintiffs' *derivative rights* in search functions, depriving Plaintiffs of revenues or other benefits they would gain from licensed search markets; (4) Google's storage of digital copies exposes Plaintiffs to the risk that hackers will make their books freely (or cheaply) available on the Internet, destroying the value of their copyrights; and

(5) Google's distribution of digital copies to participant libraries is not a transformative use, and it subjects Plaintiffs to the risk of loss of copyright revenues through access allowed by libraries. We reject these arguments and conclude that the district court correctly sustained Google's fair use defense.

Google's making of a digital copy to provide a search function is a transformative use, which augments public knowledge by making available information *about* Plaintiffs' books without providing the public with a substantial substitute for matter protected by the Plaintiffs' copyright interests in the original works or derivatives of them. The same is true, at least under present conditions, of Google's provision of the snippet function. Plaintiffs' contention that Google has usurped their opportunity to access paid and unpaid licensing markets for substantially the same functions that Google provides fails, in part because the licensing markets in fact involve very different functions than those that Google provides, and in part because an author's derivative rights do not include an exclusive right to supply information (of the sort provided by Google) about her works. Google's profit motivation does not in these circumstances justify denial of fair use. Google's program does not, at this time and on the record before us, expose Plaintiffs to an unreasonable risk of loss of copyright value through incursions of hackers. Finally, Google's provision of digital copies to participating libraries, authorizing them to make non-infringing uses, is non-infringing, and the mere speculative possibility that the libraries might allow use of their copies in an infringing manner does not make Google a contributory infringer. Plaintiffs have failed to show a material issue of fact in dispute.

We affirm the judgment.

BACKGROUND

I. Plaintiffs

The author-plaintiffs are Jim Bouton, author of *Ball Four*; Betty Miles, author of *The Trouble with Thirteen*; and Joseph Goulden, author of *The Superlawyers: The Small and Powerful World of the Great Washington Law Firms*. Each of them has a legal or beneficial ownership in the copyright for his or her book.[1] Their books have been scanned without their permission by Google, which made them available to Internet users for search and snippet view on Google's website.[2]

II. Google Books and the Google Library Project

Google's Library Project, which began in 2004, involves bi-lateral agreements between Google and a number of the world's major research libraries.[3] Under these agreements, the

[1] The Authors Guild, a membership organization of published authors, is also a plaintiff and appellant, seeking injunctive and declaratory relief on behalf of its members. However, in a separate case, this court found that, under the Copyright Act, the Authors Guild lacks standing to sue for copyright infringement on its members' behalf. *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 94 (2d Cir. 2014). As the three individual author-plaintiffs clearly do have standing, their suit and their appeal are properly adjudicated, notwithstanding the Authors Guild's lack of standing. *See Bowsher v. Synar*, 478 U.S. 714, 721 (1986) (finding that where one appellee had standing, the court need not consider the standing of other appellees in order to determine the merits of the appeal).

[2] Google now honors requests to remove books from snippet view. Some Plaintiffs appear to have had books removed from snippet view.

[3] Libraries participating in the Library Project at the time the suit was filed included the University of Michigan, the University of California, Harvard University, Stanford University,

participating libraries select books from their collections to submit to Google for inclusion in the project. Google makes a digital scan of each book, extracts a machine-readable text, and creates an index of the machine-readable text of each book. Google retains the original scanned image of each book, in part so as to improve the accuracy of the machine-readable texts and indices as image-to-text conversion technologies improve.

Since 2004, Google has scanned, rendered machine-readable, and indexed more than 20 million books, including both copyrighted works and works in the public domain. The vast majority of the books are non-fiction, and most are out of print. All of the digital information created by Google in the process is stored on servers protected by the same security systems Google uses to shield its own confidential information.

The digital corpus created by the scanning of these millions of books enables the Google Books search engine. Members of the public who access the Google Books website can enter search words or terms of their own choice, receiving in response a list of all books in the database in which those terms appear, as well as the number of times the term appears in each book. A brief description of each book, entitled "About the Book," gives some rudimentary additional information, including a list of the words and terms that appear with most frequency in the book. It sometimes provides links to buy the book online and identifies libraries where the book can be

---

Oxford University, Columbia University, Princeton University, Ghent University, Keio University, the Austrian National Library, and the New York Public Library.

found.[4] The search tool permits a researcher to identify those books, out of millions, that do, as well as those that do not, use the terms selected by the researcher. Google notes that this identifying information instantaneously supplied would otherwise not be obtainable in lifetimes of searching.

No advertising is displayed to a user of the search function. Nor does Google receive payment by reason of the searcher's use of Google's link to purchase the book.

The search engine also makes possible new forms of research, known as "text mining" and "data mining." Google's "ngrams" research tool draws on the Google Library Project corpus to furnish statistical information to Internet users about the frequency of word and phrase usage over centuries.[5] This tool permits users to discern fluctuations of interest in a particular subject over time and space by showing increases and decreases in the frequency of reference and usage in different periods and different linguistic regions. It also allows researchers to comb over the tens of millions of books Google has scanned in order to examine "word frequencies, syntactic patterns, and thematic markers" and to derive information on how nomenclature, linguistic usage, and literary style have changed over time. *Authors Guild, Inc.*, 954 F. Supp. 2d at 287. The district court gave as an example "track[ing] the frequency of references to the United States as a

[4] Appendix A exhibits, as an example, a web page that would be revealed to a searcher who entered the phase "fair use," showing snippets from ALAN LATMAN, ROBERT A. GORMAN, & JANE C. GINSBURG, COPYRIGHT FOR THE EIGHTIES (1985).

[5] Appendix B exhibits the ngram for the phrase "fair use."

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." *Id.*[6]

The Google Books search function also allows the user a limited viewing of text. In addition to telling the number of times the word or term selected by the searcher appears in the book, the search function will display a maximum of three "snippets" containing it. A snippet is a horizontal segment comprising ordinarily an eighth of a page. Each page of a conventionally formatted book[7] in the Google Books database is divided into eight non-overlapping horizontal segments, each such horizontal segment being a snippet. (Thus, for such a book with 24 lines to a page, each snippet is comprised of three lines of text.) Each search for a particular word or term within a book will reveal the same three snippets, regardless of the number of computers from which the search is launched. Only the first usage of the term on a given page is displayed. Thus, if the top snippet of a page contains two (or more) words for which the user searches, and Google's

[6] For discussions and examples of scholarship and journalism powered by searchable digital text repositories, see, e.g., David Bamman & David Smith, *Extracting Two Thousand Years of Latin from a Million Book Library*, J. COMPUTING & CULTURAL HERITAGE 5 (2012), 1-13; Jean-Baptiste Michel et al., *Quantitative Analysis of Culture Using Millions of Digitized Books*, SCIENCE 331 (Jan. 14, 2011), 176-182; Marc Egnal, *Evolution of the Novel in the United States: The Statistical Evidence*, 37 SOC. SCI. HIST. 231 (2013); Catherine Rampell, *The 'New Normal' Is Actually Pretty Old*, N.Y. TIMES ECONOMIX BLOG (Jan. 11, 2011), http://economix.blogs.nytimes.com/2011/01/11/the-new-normal-is-actually-pretty-old/?_r=0; and Christopher Forstall et al., *Modeling the Scholars: Detecting Intertextuality through Enhanced Word-Level N-Gram Matching*, DIGITAL SCHOLARSHIP IN THE HUMANITIES (May 15, 2014), http://dx.doi.org/10.1093/llc/fqu014.

[7] For unconventionally formatted books, the number of snippets per page may vary so as to approximate the same effect. The pages of a book of unusually tall, narrow format may be divided into more than eight horizontal snippets, while the pages of an unusually wide, short book may be divided into fewer than eight snippets.

program is fixed to reveal that particular snippet in response to a search for either term, the second search will duplicate the snippet already revealed by the first search, rather than moving to reveal a different snippet containing the word because the first snippet was already revealed. Google's program does not allow a searcher to increase the number of snippets revealed by repeated entry of the same search term or by entering searches from different computers. A searcher can view more than three snippets of a book by entering additional searches for different terms. However, Google makes permanently unavailable for snippet view one snippet on each page and one complete page out of every ten—a process Google calls "blacklisting."

Google also disables snippet view entirely for types of books for which a single snippet is likely to satisfy the searcher's present need for the book, such as dictionaries, cookbooks, and books of short poems. Finally, since 2005, Google will exclude any book altogether from snippet view at the request of the rights holder by the submission of an online form.

Under its contracts with the participating libraries, Google allows each library to download copies—of both the digital image and machine-readable versions—of the books that library submitted to Google for scanning (but not of books submitted by other libraries). This is done by giving each participating library access to the Google Return Interface ("GRIN"). The agreements between Google and the libraries, although not in all respects uniform, require the libraries to abide by copyright law in utilizing the digital copies they download and to take precautions to prevent dissemination of their digital copies to the public at large.[8] Through the

[8] For example, the "Cooperative Agreement" between Google and the University of Michigan ("U of M") provides, *inter alia*, that:

GRIN facility, participant libraries have downloaded at least 2.7 million digital copies of their own volumes.

> Both Google and U of M agree and intend to perform this Agreement pursuant to copyright law. If at any time, either party becomes aware of copyright infringement under this agreement, that party shall inform the other as quickly as reasonably possible. . . . U of M shall have the right to use the U of M Digital Copy . . . as part of services offered on U of M's website. U of M shall implement technological measures (e.g., through use of the robots.txt protocol) to restrict automated access to any portion of the U of M Digital Copy or the portions of the U of M website on which any portion of the U of M Digital Copy is available. U of M shall also make reasonable efforts (including but not limited to restrictions placed in Terms of Use for the U of M website) to prevent third parties from (a) downloading or otherwise obtaining any portion of the U of M Digital Copy for commercial purposes, (b) redistributing any portions of the U of M Digital Copy, or (c) automated and systematic downloading from its website image files from the U of M Digital Copy. U of M shall restrict access to the U of M Digital Copy to those persons having a need to access such materials and shall also cooperate in good faith with Google to mutually develop methods and systems for ensuring that the substantial portions of the U of M Digital Copy are not downloaded from the services offered on U of M's website or otherwise disseminated to the public at large.

JA 233.

Google's agreement with Stanford appears to be less restrictive on Stanford than its agreements with other libraries. It ostensibly permits Stanford's libraries to "provide access to or copies from the Stanford Digital Copy" to a wide range of users, including individuals authorized to access the Stanford University Network, individuals affiliated with "partner research libraries," and "education, research, government institutions and libraries not affiliated with Stanford," CA 133, and to permit authorized individuals to download or print up to ten percent of Stanford Digital Copy. On the other hand, the agreement requires Stanford to employ its digital copies in conformity with the copyright law. Without evidence to the contrary, which Plaintiffs have not provided, it seems reasonable to construe these potentially conflicting provisions as meaning that Stanford may do the enumerated things ostensibly permitted only to the extent that doing so would be in conformity with the copyright law.

III. Procedural History

Plaintiffs brought this suit on September 20, 2005, as a putative class action on behalf of similarly situated, rights-owning authors.[9] After several years of negotiation, the parties reached a proposed settlement that would have resolved the claims on a class-wide basis. The proposed settlement allowed Google to make substantially more extensive use of its scans of copyrighted books than contemplated under the present judgment, and provided that Google would make payments to the rights holders in return. On March 22, 2011, however, the district court rejected the proposed settlement as unfair to the class members who relied on the named plaintiffs to represent their interests. *Authors Guild v. Google Inc.*, 770 F. Supp. 2d 666, 679-680 (S.D.N.Y. 2011).

On October 14, 2011, Plaintiffs filed a fourth amended class action complaint, which is the operative complaint for this appeal. *See* Dist. Ct. Docket No. 985. The district court certified a class on May 31, 2012. *Authors Guild v. Google Inc.*, 282 F.R.D. 384 (S.D.N.Y. 2012). Google appealed from the certification, and moved in the district court for summary judgment on its fair use defense. Plaintiffs cross-moved in the district court for summary judgment. On the appeal from the class certification, our court—questioning whether it was reasonable to infer that the putative class of authors favored the relief sought by the named plaintiffs—provisionally vacated that class certification without addressing the merits of the issue, concluding instead that "resolution of Google's fair use defense in the first instance will necessarily inform and perhaps

[9] A year earlier, authors brought suit against the HathiTrust Digital Library, alleging facts that are closely related, although not identical, to those alleged in the instant case. *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 91 (2d Cir. 2014).

moot our analysis of many class certification issues." *Authors Guild, Inc. v. Google Inc.*, 721 F.3d 132, 134 (2d Cir. 2013).

On November 14, 2013, the district court granted Google's motion for summary judgment, concluding that the uses made by Google of copyrighted books were fair uses, protected by § 107. *Authors Guild*, 954 F. Supp. 2d at 284. Upon consideration of the four statutory factors of § 107, the district court found that Google's uses were transformative, that its display of copyrighted material was properly limited, and that the Google Books program did not impermissibly serve as a market substitute for the original works. *Id.* at 290. The court entered judgment initially on November 27, 2013, followed by an amended judgment on December 10, 2013, dismissing Plaintiffs' claims with prejudice. Plaintiffs filed timely notice of appeal.

DISCUSSION[10]

I. The Law of Fair Use

The ultimate goal of copyright is to expand public knowledge and understanding, which copyright seeks to achieve by giving potential creators exclusive control over copying of their works, thus giving them a financial incentive to create informative, intellectually enriching works for public consumption. This objective is clearly reflected in the Constitution's empowerment of Congress "*To promote the Progress of Science* . . . by securing for limited

[10] The district court had subject-matter jurisdiction over this federal copyright action pursuant to 28 U.S.C. §§ 1331 and 1338(a). This court has jurisdiction over the appeal from the final decision of the district court pursuant to 28 U.S.C. § 1291. We review an order granting summary judgment *de novo*, drawing all reasonable factual inferences in favor of the non-moving party. *Ment Bros. Iron Works Co., Inc. v. Interstate Fire & Cas. Co.*, 702 F.3d 118, 120–21 (2d Cir. 2012).

Times to Authors . . . the exclusive Right to their respective Writings." U.S. Const., Art. I, § 8, cl. 8) (emphasis added).[11] Thus, while authors are undoubtedly important intended beneficiaries of copyright, the ultimate, primary intended beneficiary is the public, whose access to knowledge copyright seeks to advance by providing rewards for authorship.

For nearly three hundred years, since shortly after the birth of copyright in England in 1710,[12] courts have recognized that, in certain circumstances, giving authors *absolute* control over all copying from their works would tend in some circumstances to limit, rather than expand, public knowledge. In the words of Lord Ellenborough, "[W]hile I shall think myself bound to secure every man in the enjoyment of his copy-right, one must not put manacles upon science." *Cary v. Kearsley*, 170 Eng. Rep. 679, 681, 4 Esp. 168, 170 (1802). Courts thus developed the doctrine, eventually named fair use, which permits unauthorized copying in some circumstances, so as to further "copyright's very purpose, '[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). Although well established in the common law development of copyright, fair use was not recognized in the terms of our statute until the adoption of § 107 in the Copyright Act of 1976. 17 U.S.C. §§ 101 *et seq.*

---

[11] A similar message is reflected in England's original copyright enactment, "An Act for the Encouragement of Learning, by Vesting the Copies of Printed Books in the Authors," which explains as its purpose "the Encouragement of Learned Men to Compose and Write useful Books." Statute of Anne [1710].

[12] *Id.*

Section 107, in its present form,[13] provides:

> [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—
>
> (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
> (2) the nature of the copyrighted work;
> (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
> (4) the effect of the use upon the potential market for or value of the copyrighted work.
>
> The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.

17 U.S.C. § 107. As the Supreme Court has designated fair use an affirmative defense, *see* *Campbell*, 510 U.S. at 590, the party asserting fair use bears the burden of proof, *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 918 (2d Cir. 1994).

The statute's wording, derived from a brief observation of Justice Joseph Story in *Folsom v. Marsh*,[14] does not furnish standards for recognition of fair use. Its instruction to consider the "purpose and character" of the secondary use and the "nature" of the copyrighted work does not explain what types of "purpose and character" or "nature" favor a finding of fair use and which do not. In fact, as the Supreme Court observed in *Campbell*, the House Report makes clear that, in

---

[13] The last sentence was added in 1992, rejecting this court's assertion in *Salinger v. Random House, Inc.* that unpublished works "normally enjoy insulation from fair use copying." 811 F.2d 90, 95 (2d Cir. 1987). *See* Pub. L. 102–492, Oct. 24, 1992, 106 Stat. 3145.

[14] 9 F. Cas. 342, 348 (C.C.D. Mass. 1841) ("[W]e must often, in deciding questions of this sort, look to the nature and objects of the selections made, the quantity and value of the materials used, and the degree in which the use may prejudice the sale, or diminish the profits, or supersede the objects, of the original work.").

passing the statute, Congress had no intention of normatively dictating fair use policy. The purpose of the enactment was to give recognition in the statute itself to such an important part of copyright law developed by the courts through the common law process. "Congress meant § 107 'to restate the present judicial doctrine of fair use, not to change, narrow, or enlarge it an any way,' and intended that courts continue the common-law tradition of fair use adjudication." *Campbell*, 510 U.S. at 577 (quoting H.R. Rep. No. 94-1476, at 66 (1976), S. Rep. No. 94-473, at 62 (1975), U.S. Code Cong. & Admin. News 5659, 5679 (1976)). Furthermore, notwithstanding fair use's long common-law history, not until the *Campbell* ruling in 1994 did courts undertake to explain the standards for finding fair use.

The *Campbell* Court undertook a comprehensive analysis of fair use's requirements, discussing every segment of § 107. Beginning with the examples of purposes set forth in the statute's preamble, the Court made clear that they are "illustrative and not limitative" and "provide only general guidance about the sorts of copying that courts and Congress most commonly ha[ve] found to be fair uses." 510 U.S. at 577-578 (internal quotations and citations omitted). The statute "calls for case-by-case analysis" and "is not to be simplified with bright-line rules." *Id.* at 577. Section 107's four factors are not to "be treated in isolation, one from another. All are to be explored, and the results weighed together, in light of the purposes of copyright." *Id.* at 578. Each factor thus stands as part of a multifaceted assessment of the crucial question: how to define the boundary limit of the original author's exclusive rights in order to best serve the overall objectives of the copyright law to expand public learning while protecting the incentives of authors to create for the public good.

At the same time, the Supreme Court has made clear that some of the statute's four listed factors are more significant than others. The Court observed in *Harper & Row Publishers, Inc. v. Nation Enterprises* that the fourth factor, which assesses the harm the secondary use can cause to the market for, or the value of, the copyright for the original, "is undoubtedly the single most important element of fair use." 471 U.S. 539, 566 (1985) (citing MELVILLE B. NIMMER, 3 NIMMER ON COPYRIGHT § 13.05[A], at 13–76 (1984)). This is consistent with the fact that the copyright is a commercial right, intended to protect the ability of authors to profit from the exclusive right to merchandise their own work.

In *Campbell*, the Court stressed also the importance of the first factor, the "purpose and character of the secondary use." 17 U.S.C. § 107(1). The more the appropriator is using the copied material for new, transformative purposes, the more it serves copyright's goal of enriching public knowledge and the less likely it is that the appropriation will serve as a substitute for the original or its plausible derivatives, shrinking the protected market opportunities of the copyrighted work. 510 U.S. at 591 (noting that, when the secondary use is transformative, "market substitution is at least less certain, and market harm may not be so readily inferred.").

With this background, we proceed to discuss each of the statutory factors, as illuminated by *Campbell* and subsequent case law, in relation to the issues here in dispute.

II. The Search and Snippet View Functions

A. Factor One

(1) Transformative purpose. *Campbell*'s explanation of the first factor's inquiry into the "purpose and character" of the secondary use focuses on whether the new work, "in Justice Story's

words, . . . merely 'supersede[s] the objects' of the original creation, . . . or instead adds something new, with a further purpose . . . . [I]t asks, in other words, whether and to what extent the new work is 'transformative.'" 510 U.S. at 578-579 (citations omitted). While recognizing that a transformative use is "not absolutely necessary for a finding of fair use," the opinion further explains that the "goal of copyright, to promote science and the arts, is generally furthered by the creation of transformative works" and that "[s]uch works thus lie at the heart of the fair use doctrine's guarantee of breathing space within the confines of copyright." *Id.* at 579. In other words, transformative uses tend to favor a fair use finding because a transformative use is one that communicates something new and different from the original or expands its utility, thus serving copyright's overall objective of contributing to public knowledge.

The word "transformative" cannot be taken too literally as a sufficient key to understanding the elements of fair use. It is rather a suggestive symbol for a complex thought, and does not mean that any and all changes made to an author's original text will necessarily support a finding of fair use. The Supreme Court's discussion in *Campbell* gave important guidance on assessing when a transformative use tends to support a conclusion of fair use. The defendant in that case defended on the ground that its work was a parody of the original and that parody is a time-honored category of fair use. Explaining why parody makes a stronger, or in any event more obvious, claim of fair use than satire, the Court stated,

> [T]he heart of any parodist's claim to quote from existing material . . . is the use of . . . a prior author's composition to . . .*comment[] on that author's works*. . . . If, on the contrary, the commentary has no critical bearing on the substance or style of the original composition, which the alleged infringer merely uses to get attention or to avoid the drudgery in working up something fresh, the claim to fairness in borrowing from another's work diminishes accordingly (if it does not vanish). . . . Parody needs to mimic an original

> to make its point, and so has some claim to use the creation of its victim's . . . imagination, whereas satire can stand on its own two feet and so requires justification for the very act of borrowing.

Id. at 580-81 (emphasis added). In other words, the would-be fair user of another's work must have justification for the taking. A secondary author is not necessarily at liberty to make wholesale takings of the original author's expression merely because of how well the original author's expression would convey the secondary author's different message. Among the best recognized justifications for copying from another's work is to provide comment on it or criticism of it. A taking from another author's work for the purpose of making points that have no bearing on the original may well be fair use, but the taker would need to show a justification. This part of the Supreme Court's discussion is significant in assessing Google's claim of fair use because, as discussed extensively below, Google's claim of transformative purpose for copying from the works of others is to provide otherwise unavailable information about the originals.

A further complication that can result from oversimplified reliance on whether the copying involves transformation is that the word "transform" also plays a role in defining "derivative works," over which the original rights holder retains exclusive control. Section 106 of the Act specifies the "exclusive right[]" of the copyright owner "(2) to prepare derivative works based upon the copyrighted work." *See* 17 U.S.C. § 106. The statute defines derivative works largely by example, rather than explanation. The examples include "translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgement, condensation," to which list the statute adds "any other form in which a work may

be . . . *transformed*." 17 U.S.C. § 101 (emphasis added).[15] As we noted in *Authors Guild, Inc. v. HathiTrust*, "[p]aradigmatic examples of derivative works include the translation of a novel into another language, the adaptation of a novel into a movie or play, or the recasting of a novel as an e-book or an audiobook." 755 F.3d 87, 95 (2d Cir. 2014). While such changes can be described as transformations, they do not involve the kind of transformative purpose that favors a fair use finding. The statutory definition suggests that derivative works generally involve transformations in the nature of *changes of form*. 17 U.S.C. § 101. By contrast, copying from an original for the purpose of criticism or commentary on the original[16] or provision of information about it,[17]

[15] The full text of the statutory definition is as follows: "A 'derivative work' is a work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgement, condensation, or any other form in which a work may be recast, transformed, or adapted. A work consisting of editorial revisions, annotations, elaborations, or other modifications which, as a whole, represent an original work of authorship, is a 'derivative work.'" 17 U.S.C. § 101.

[16] *See, e.g.*, *Suntrust Bank v. Houghton Mifflin Co.*, 268 F.3d 1257, 1269-1271 (11th Cir. 2001) (copying from original to support parodic criticism of original's moral code justified as transformative fair use purpose).

[17] *See, e.g.*, *HathiTrust*, 755 F.3d at 97-98 (justifying as transformative fair use purpose the digital copying of original for purpose of permitting searchers to determine whether its text employs particular words); *A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630, 638-640 (4th Cir. 2009) (justifying as transformative fair use purpose the complete digital copying of a manuscript to determine whether the original included matter plagiarized from other works); *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1165 (9th Cir. 2007) (justifying as transformative fair use purpose the use of a digital, thumbnail copy of the original to provide an Internet pathway to the original); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818-819 (9th Cir. 2003) (same); *Bond v. Blum*, 317 F.3d 385 (4th Cir. 2003) (justifying as fair use purpose the copying of author's original unpublished autobiographical manuscript for the purpose of showing that he murdered his father and was an unfit custodian of his children); *Nuñez v. Carribean Int'l News Corp.*, 235 F.3d 18, 21-23 (1st Cir. 2000) (justifying as transformative fair

tends most clearly to satisfy *Campbell*'s notion of the "transformative" purpose involved in the analysis of Factor One.[18]

With these considerations in mind, we first consider whether Google's search and snippet views functions satisfy the first fair use factor with respect to Plaintiffs' rights in their books.

---

use purpose a newspaper's copying of a photo of winner of beauty pageant in a revealing pose for the purpose of informing the public of the reason the winner's title was withdrawn).

[18] The Seventh Circuit takes the position that the kind of secondary use that favors satisfaction of the fair use test is better described as a "complementary" use, referring to how a hammer and nail complement one another in that together they achieve results that neither can accomplish on its own. *Ty, Inc. v Publ'ns Int'l, Ltd.*, 292 F.3d 512, 517–518 (7th Cir. 2002); *see also Kienitz v. Sconnie Nation LLC*, 766 F.3d 756, 758 (7th Cir. 2014), *cert. denied*, 135 S. Ct. 1555 (2015); WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW 153–154 (2003). We do not find the term "complementary" particularly helpful in explaining fair use. The term would encompass changes of form that are generally understood to produce derivative works, rather than fair uses, and, at the same time, would fail to encompass copying for purposes that are generally and properly viewed as creating fair uses. When a novel is converted into film, for example, the original novel and the film ideally complement one another in that each contributes to achieving results that neither can accomplish on its own. The invention of the original author combines with the cinematographic interpretive skills of the filmmaker to produce something that neither could have produced independently. Nonetheless, at least when the intention of the film is to make a "motion picture version" of the novel, 17 U.S.C. § 101, without undertaking to parody it or to comment on it, the film is generally understood to be a derivative work, which under § 106, falls within the exclusive rights of the copyright owner. Although they complement one another, the film is not a fair use. At the same time, when a secondary work quotes an original for the purpose of parodying it, or discrediting it by exposing its inaccuracies, illogic, or dishonesty, such an undertaking is not within the exclusive prerogatives of the rights holder; it produces a fair use. Yet, when the purpose of the second is essentially to destroy the first, the two are not comfortably described as complementaries that combine to produce together something that neither could have produced independently of the other. We recognize, as just noted above, that the word "transformative," if interpreted too broadly, can also seem to authorize copying that should fall within the scope of an author's derivative rights. Attempts to find a circumspect shorthand for a complex concept are best understood as suggestive of a general direction, rather than as definitive descriptions.

(The question whether these functions might infringe upon Plaintiffs' derivative rights is discussed in the next Part.)

(2) Search Function. We have no difficulty concluding that Google's making of a digital copy of Plaintiffs' books for the purpose of enabling a search for identification of books containing a term of interest to the searcher involves a highly transformative purpose, in the sense intended by *Campbell*. Our court's exemplary discussion in *HathiTrust* informs our ruling. That case involved a dispute that is closely related, although not identical, to this one. Authors brought claims of copyright infringement against HathiTrust, an entity formed by libraries participating in the Google Library Project to pool the digital copies of their books created for them by Google. The suit challenged various usages HathiTrust made of the digital copies. Among the challenged uses was HathiTrust's offer to its patrons of "full-text searches," which, very much like the search offered by Google Books to Internet users, permitted patrons of the libraries to locate in which of the digitized books specific words or phrases appeared. 755 F.3d at 98. (HathiTrust's search facility did not include the snippet view function, or any other display of text.) We concluded that both the making of the digital copies and the use of those copies to offer the search tool were fair uses. *Id.* at 105.

Notwithstanding that the libraries had downloaded and stored complete digital copies of entire books, we noted that such copying was essential to permit searchers to identify and locate the books in which words or phrases of interest to them appeared. *Id.* at 97. We concluded "that the creation of a full-text searchable database is a quintessentially transformative use . . . [as] the result of a word search is different in purpose, character, expression, meaning, and message from

the page (and the book) from which it is drawn." *Id.* We cited *A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630, 639-40 (4th Cir. 2009), *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1165 (9th Cir. 2007), and *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 819 (9th Cir. 2003) as examples of cases in which courts had similarly found the creation of complete digital copies of copyrighted works to be transformative fair uses when the copies "served a different function from the original." *HathiTrust*, 755 F.3d at 97.

As with *HathiTrust* (and *iParadigms*), the purpose of Google's copying of the original copyrighted books is to make available significant information *about those books*, permitting a searcher to identify those that contain a word or term of interest, as well as those that do not include reference to it. In addition, through the ngrams tool, Google allows readers to learn the frequency of usage of selected words in the aggregate corpus of published books in different historical periods. We have no doubt that the purpose of this copying is the sort of transformative purpose described in *Campbell* as strongly favoring satisfaction of the first factor.

We recognize that our case differs from *HathiTrust* in two potentially significant respects. First, HathiTrust did not "display to the user any text from the underlying copyrighted work," 755 F.3d at 91, whereas Google Books provides the searcher with snippets containing the word that is the subject of the search. Second, HathiTrust was a nonprofit educational entity, while Google is a profit-motivated commercial corporation. We discuss those differences below.

(3) Snippet View. Plaintiffs correctly point out that this case is significantly different from *HathiTrust* in that the Google Books search function allows searchers to read snippets from the book searched, whereas HathiTrust did not allow searchers to view any part of the book.

Snippet view adds important value to the basic transformative search function, which tells only whether and how often the searched term appears in the book. Merely knowing that a term of interest appears in a book does not necessarily tell the searcher whether she needs to obtain the book, because it does not reveal whether the term is discussed in a manner or context falling within the scope of the searcher's interest. For example, a searcher seeking books that explore Einstein's theories, who finds that a particular book includes 39 usages of "Einstein," will nonetheless conclude she can skip that book if the snippets reveal that the book speaks of "Einstein" because that is the name of the author's cat. In contrast, the snippet will tell the searcher that this is a book she needs to obtain if the snippet shows that the author is engaging with Einstein's theories.

Google's division of the page into tiny snippets is designed to show the searcher just enough context surrounding the searched term to help her evaluate whether the book falls within the scope of her interest (without revealing so much as to threaten the author's copyright interests). Snippet view thus adds importantly to the highly transformative purpose of identifying books of interest to the searcher. With respect to the first factor test, it favors a finding of fair use (unless the value of its transformative purpose is overcome by its providing text in a manner that offers a competing substitute for Plaintiffs' books, which we discuss under factors three and four below).

(4) Google's Commercial Motivation. Plaintiffs also contend that Google's commercial motivation weighs in their favor under the first factor. Google's commercial motivation distinguishes this case from *HathiTrust*, as the defendant in that case was a non-profit entity

founded by, and acting as the representative of, libraries. Although Google has no revenues flowing directly from its operation of the Google Books functions, Plaintiffs stress that Google is profit-motivated and seeks to use its dominance of book search to fortify its overall dominance of the Internet search market, and that thereby Google indirectly reaps profits from the Google Books functions.

For these arguments Plaintiffs rely primarily on two sources. First is Congress's specification in spelling out the first fair use factor in the text of § 107 that consideration of the "purpose and character of the [secondary] use" should "include[e] whether such use is of a commercial nature or is for nonprofit educational purposes." Second is the Supreme Court's assertion in dictum in *Sony Corporation of America v. Universal City Studios, Inc*, that "every commercial use of copyrighted material is presumptively . . . unfair." 464 U.S. 417, 451 (1984). If that were the extent of precedential authority on the relevance of commercial motivation, Plaintiffs' arguments would muster impressive support. However, while the commercial motivation of the secondary use can undoubtedly weigh against a finding of fair use in some circumstances, the Supreme Court, our court, and others have eventually recognized that the *Sony* dictum was enormously overstated.[19]

[19] *Campbell*, 510 U.S. at 583–84; *Cariou v. Prince*, 714 F.3d 694, 708 (2d Cir. 2013) *cert. denied*, 134 S. Ct. 618 (2013); *Castle Rock Entm't, Inc. v. Carol Pub. Grp., Inc.*, 150 F.3d 132, 141-42 (2d Cir. 1998); *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1165 (9th Cir. 2007); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 819 (9th Cir. 2003); *see also Monge v. Maya Magazines, Inc.*, 688 F.3d 1164, 1172 (9th Cir. 2012) (noting that *Campbell* "debunked the notion that *Sony* called for a 'hard evidentiary presumption' that commercial use is presumptively unfair.")

The Sixth Circuit took the *Sony* dictum at its word in *Acuff-Rose Music, Inc. v. Campbell*, concluding that, because the defendant rap music group's spoof of the plaintiff's ballad was done for profit, it could not be fair use. 972 F.2d 1429, 1436-1437 (6th Cir. 1992). The Supreme Court reversed on this very point, observing that "Congress could not have intended" such a broad presumption against commercial fair uses, as "nearly all of the illustrative uses listed in the preamble paragraph of § 107 . . . are generally conducted for profit in this country." *Campbell*, 510 U.S. at 584 (internal quotation marks and citations omitted). The Court emphasized Congress's statement in the House Report to the effect that the commercial or nonprofit character of a work is "not conclusive" but merely "a fact to be 'weighed along with other[s] in fair use decisions." *Id.* at 585 (quoting H.R. Rep. No. 94-1476, at 66 (1976)). In explaining the first fair use factor, the Court clarified that "the more transformative the [secondary] work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use." *Id.* at 579.

Our court has since repeatedly rejected the contention that commercial motivation should outweigh a convincing transformative purpose and absence of significant substitutive competition with the original. *See Cariou v. Prince*, 714 F.3d 694, 708 (2d Cir. 2013), *cert. denied*, 134 S. Ct. 618 (2013) ("The commercial/nonprofit dichotomy concerns the unfairness that arises when a secondary user makes unauthorized use of copyrighted material to capture significant revenues as a direct consequence of copying the original work. This factor must be applied with caution because, as the Supreme Court has recognized, Congress could not have intended a rule that commercial uses are presumptively unfair. Instead, the more transformative

the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use.") (internal quotation marks, citations, and alterations omitted); *Castle Rock Entm't, Inc. v. Carol Pub. Grp., Inc.*, 150 F.3d 132, 141-42 (2d Cir. 1998) ("We . . . do not give much weight to the fact that the secondary use was for commercial gain. The more critical inquiry under the first factor and in fair use analysis generally is whether the allegedly infringing work merely supersedes the original work or instead adds something new, with a further purpose or different character, altering the first with new meaning or message, in other words whether and to what extent the new work is transformative.") (internal quotation marks, citations, and alterations omitted).

While we recognize that in some circumstances, a commercial motivation on the part of the secondary user will weigh against her, especially, as the Supreme Court suggested, when a persuasive transformative purpose is lacking, *Campbell*, 510 U.S. at 579, we see no reason in this case why Google's overall profit motivation should prevail as a reason for denying fair use over its highly convincing transformative purpose, together with the absence of significant substitutive competition, as reasons for granting fair use. Many of the most universally accepted forms of fair use, such as news reporting and commentary, quotation in historical or analytic books, reviews of books, and performances, as well as parody, are all normally done commercially for profit.[20]

[20] Just as there is no reason for presuming that a commercial use is not a fair use, which would defeat the most widely accepted and logically justified areas of fair use, there is likewise no reason to presume categorically that a nonprofit educational purpose should qualify as a fair use. Authors who write for educational purposes, and publishers who invest substantial funds to

B. Factor Two

The second fair use factor directs consideration of the "nature of the copyrighted work." While the "transformative purpose" inquiry discussed above is conventionally treated as a part of first factor analysis, it inevitably involves the second factor as well. One cannot assess whether the copying work has an objective that differs from the original without considering both works, and their respective objectives.

The second factor has rarely played a significant role in the determination of a fair use dispute. *See* WILLIAM F. PATRY, PATRY ON FAIR USE § 4.1 (2015). The Supreme Court in *Harper & Row* made a passing observation in dictum that, "[t]he law generally recognizes a greater need to disseminate factual works than works of fiction or fantasy." 471 U.S. 539, 563 (1985). Courts have sometimes speculated that this might mean that a finding of fair use is more favored when the copying is of factual works than when copying is from works of fiction. However, while the copyright does not protect facts or ideas set forth in a work, it does protect that author's manner of expressing those facts and ideas. At least unless a persuasive fair use justification is involved, authors of factual works, like authors of fiction, should be entitled to copyright protection of their protected expression. The mere fact that the original is a factual work therefore should not imply that others may freely copy it. Those who report the news undoubtedly create factual works. It

---

publish educational materials, would lose the ability to earn revenues if users were permitted to copy the materials freely merely because such copying was in the service of a nonprofit educational mission. The publication of educational materials would be substantially curtailed if such publications could be freely copied for nonprofit educational purposes.

cannot seriously be argued that, for that reason, others may freely copy and re-disseminate news reports.[21]

In considering the second factor in *HathiTrust*, we concluded that it was "not dispositive," 755 F.3d at 98, commenting that courts have hardly ever found that the second factor in isolation played a large role in explaining a fair use decision. The same is true here. While each of the three Plaintiffs' books in this case is factual, we do not consider that as a boost to Google's claim of fair use. If one (or all) of the plaintiff works were fiction, we do not think that would change in any way our appraisal. Nothing in this case influences us one way or the other with respect to the second factor considered in isolation. To the extent that the "nature" of the original copyrighted work necessarily combines with the "purpose and character" of the secondary work to permit assessment of whether the secondary work uses the original in a "transformative" manner, as the term is used in *Campbell*, the second factor favors fair use not because Plaintiffs' works are factual, but because the secondary use transformatively provides valuable information about the original, rather than replicating protected expression in a manner that provides a meaningful substitute for the original.

[21] We think it unlikely that the Supreme Court meant in its concise dictum that secondary authors are at liberty to copy extensively from the protected expression of the original author merely because the material is factual. What the *Harper & Row* dictum may well have meant is that, because in the case of factual writings, there is often occasion to test the accuracy of, to rely on, or to repeat their factual propositions, and such testing and reliance may reasonably require quotation (lest a change of expression unwittingly alter the facts), factual works often present well justified fair uses, even if the mere fact that the work is factual does not necessarily justify copying of its protected expression.

C. Factor Three

The third statutory factor instructs us to consider "the amount and substantiality of the portion used in relation to the copyrighted work as a whole." The clear implication of the third factor is that a finding of fair use is more likely when small amounts, or less important passages, are copied than when the copying is extensive, or encompasses the most important parts of the original.[22] The obvious reason for this lies in the relationship between the third and the fourth factors. The larger the amount, or the more important the part, of the original that is copied, the greater the likelihood that the secondary work might serve as an effectively competing substitute for the original, and might therefore diminish the original rights holder's sales and profits.

(1) Search Function. The Google Books program has made a digital copy of the entirety of each of Plaintiffs' books. Notwithstanding the reasonable implication of Factor Three that fair use is more likely to be favored by the copying of smaller, rather than larger, portions of the original, courts have rejected any categorical rule that a copying of the entirety cannot be a fair use.[23] Complete unchanged copying has repeatedly been found justified as fair use when the copying was reasonably appropriate to achieve the copier's transformative purpose and was done in such a

[22] *See Harper & Row*, 471 U.S. at 564-565 (rejecting fair use defense for copying of only about 300 words, where the portion copied was deemed "the heart of the book").

[23] Some copyright scholars have argued this position. *See, e.g.*, Paul Goldstein, *Copyright's Commons*, 29 COLUM. J.L. & ARTS 1, 5-6 (2005).

manner that it did not offer a competing substitute for the original.[24] The Supreme Court said in *Campbell* that "the extent of permissible copying varies with the purpose and character of the use" and characterized the relevant questions as whether "the amount and substantiality of the portion used . . . are reasonable in relation to the purpose of the copying," *Campbell*, 510 U.S. at 586-587, noting that the answer to that question will be affected by "the degree to which the [copying work] may serve as a market substitute for the original or potentially licensed derivatives," *id.* at 587-588 (finding that, in the case of a parodic song, "how much . . . is reasonable will depend, say, on the extent to which the song's overriding purpose and character is to parody the original or, in contrast, the likelihood that the parody may serve as a market substitute for the original").

In *HathiTrust*, our court concluded in its discussion of the third factor that "[b]ecause it was reasonably necessary for the [HathiTrust Digital Library] to make use of the entirety of the works in order to enable the full-text search function, we do not believe the copying was excessive." 755 F.3d at 98. As with *HathiTrust*, not only is the copying of the totality of the original reasonably appropriate to Google's transformative purpose, it is literally necessary to achieve that purpose. If Google copied less than the totality of the originals, its search function could not advise searchers reliably whether their searched term appears in a book (or how many times).

While Google *makes* an unauthorized digital copy of the entire book, it does not reveal that digital copy to the public. The copy is made to enable the search functions to reveal limited,

[24] *See* cases cited *supra* note 17; *see also Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 613 (2d Cir. 2006) ("[C]opying the entirety of a work is sometimes necessary to make a fair use of the [work].").

important information about the books. With respect to the search function, Google satisfies the third factor test, as illuminated by the Supreme Court in *Campbell*.

(2) Snippet View. Google's provision of snippet view makes our third factor inquiry different from that inquiry in *HathiTrust*. What matters in such cases is not so much "the amount and substantiality of the portion used" *in making a copy*, but rather the amount and substantiality of *what is thereby made accessible* to a public for which it may serve as a competing substitute. In *HathiTrust*, notwithstanding the defendant's full-text copying, the search function revealed virtually nothing of the text of the originals to the public. Here, through the snippet view, more is revealed to searchers than in *HathiTrust*.

Without doubt, enabling searchers to see portions of the copied texts could have determinative effect on the fair use analysis. The larger the quantity of the copyrighted text the searcher can see and the more control the searcher can exercise over what part of the text she sees, the greater the likelihood that those revelations could serve her as an effective, free substitute for the purchase of the plaintiff's book. We nonetheless conclude that, at least as presently structured by Google, the snippet view does not reveal matter that offers the marketplace a significantly competing substitute for the copyrighted work.

Google has constructed the snippet feature in a manner that substantially protects against its serving as an effectively competing substitute for Plaintiffs' books. In the Background section of this opinion, we describe a variety of limitations Google imposes on the snippet function. These include the small size of the snippets (normally one eighth of a page), the blacklisting of one snippet per page and of one page in every ten, the fact that no more than three snippets are

shown—and no more than one per page—for each term searched, and the fact that the same snippets are shown for a searched term no matter how many times, or from how many different computers, the term is searched. In addition, Google does not provide snippet view for types of books, such as dictionaries and cookbooks, for which viewing a small segment is likely to satisfy the searcher's need. The result of these restrictions is, so far as the record demonstrates, that a searcher cannot succeed, even after long extended effort to multiply what can be revealed, in revealing through a snippet search what could usefully serve as a competing substitute for the original.

The blacklisting, which permanently blocks about 22% of a book's text from snippet view, is by no means the most important of the obstacles Google has designed. While it is true that the blacklisting of 22% leaves 78% of a book *theoretically* accessible to a searcher, it does not follow that any large part of that 78% is in fact accessible. The other restrictions built into the program work together to ensure that, even after protracted effort over a substantial period of time, only small and randomly scattered portions of a book will be accessible. In an effort to show what large portions of text searchers can read through persistently augmented snippet searches, Plaintiffs' counsel employed researchers over a period of weeks to do multiple word searches on Plaintiffs' books. In no case were they able to access as much as 16% of the text, and the snippets collected were usually not sequential but scattered randomly throughout the book. Because Google's snippets are arbitrarily and uniformly divided by lines of text, and not by complete sentences, paragraphs, or any measure dictated by content, a searcher would have great difficulty constructing a search so as to provide any extensive information about the book's use

of that term. As snippet view never reveals more than one snippet per page in response to repeated searches for the same term, it is at least difficult, and often impossible, for a searcher to gain access to more than a single snippet's worth of an extended, continuous discussion of the term.

The fact that Plaintiffs' searchers managed to reveal nearly 16% of the text of Plaintiffs' books overstates the degree to which snippet view can provide a meaningful substitute. At least as important as the percentage of words of a book that are revealed is the manner and order in which they are revealed. Even if the search function revealed 100% of the words of the copyrighted book, this would be of little substitutive value if the words were revealed in alphabetical order, or any order other than the order they follow in the original book. It cannot be said that a revelation is "substantial" in the sense intended by the statute's third factor if the revelation is in a form that communicates little of the sense of the original. The fragmentary and scattered nature of the snippets revealed, even after a determined, assiduous, time-consuming search, results in a revelation that is not "substantial," even if it includes an aggregate 16% of the text of the book. If snippet view could be used to reveal a coherent block amounting to 16% of a book, that would raise a very different question beyond the scope of our inquiry.

D. Factor Four

The fourth fair use factor, "the effect of the [copying] use upon the potential market for or value of the copyrighted work," focuses on whether the copy brings to the marketplace a competing substitute for the original, or its derivative, so as to deprive the rights holder of significant revenues because of the likelihood that potential purchasers may opt to acquire the

copy in preference to the original. Because copyright is a commercial doctrine whose objective is to stimulate creativity among potential authors by enabling them to earn money from their creations, the fourth factor is of great importance in making a fair use assessment. *See Harper & Row*, 471 U.S. at 566 (describing the fourth factor as "undoubtedly the single most important element of fair use").

Campbell stressed the close linkage between the first and fourth factors, in that the more the copying is done to achieve a purpose that differs from the purpose of the original, the less likely it is that the copy will serve as a satisfactory substitute for the original. 510 U.S. at 591. Consistent with that observation, the *HathiTrust* court found that the fourth factor favored the defendant and supported a finding of fair use because the ability to search the text of the book to determine whether it includes selected words "does not serve as a substitute for the books that are being searched." 755 F.3d at 100.

However, *Campbell*'s observation as to the likelihood of a secondary use serving as an effective substitute goes only so far. Even if the *purpose* of the copying is for a valuably transformative purpose, such copying might nonetheless harm the value of the copyrighted original if done in a manner that results in widespread revelation of sufficiently significant portions of the original as to make available a significantly competing substitute. The question for us is whether snippet view, notwithstanding its transformative purpose, does that. We conclude that, at least as snippet view is presently constructed, it does not.

Especially in view of the fact that the normal purchase price of a book is relatively low in relation to the cost of manpower needed to secure an arbitrary assortment of randomly scattered

snippets, we conclude that the snippet function does not give searchers access to effectively competing substitutes. Snippet view, at best and after a large commitment of manpower, produces discontinuous, tiny fragments, amounting in the aggregate to no more than 16% of a book. This does not threaten the rights holders with any significant harm to the value of their copyrights or diminish their harvest of copyright revenue.

We recognize that the snippet function can cause *some* loss of sales. There are surely instances in which a searcher's need for access to a text will be satisfied by the snippet view, resulting in either the loss of a sale to that searcher, or reduction of demand on libraries for that title, which might have resulted in libraries purchasing additional copies. But the possibility, or even the probability or certainty, of some loss of sales does not suffice to make the copy an effectively competing substitute that would tilt the weighty fourth factor in favor of the rights holder in the original. There must be a meaningful or significant effect "upon the potential market for or value of the copyrighted work." 17 U.S.C. § 107(4).

Furthermore, the type of loss of sale envisioned above will generally occur in relation to interests that are not protected by the copyright. A snippet's capacity to satisfy a searcher's need for access to a copyrighted book will at times be because the snippet conveys a historical fact that the searcher needs to ascertain. For example, a student writing a paper on Franklin D. Roosevelt might need to learn the year Roosevelt was stricken with polio. By entering "Roosevelt polio" in a Google Books search, the student would be taken to (among numerous sites) a snippet from page 31 of Richard Thayer Goldberg's *The Making of Franklin D. Roosevelt* (1981), telling that the polio attack occurred in 1921. This would satisfy the searcher's

need for the book, eliminating any need to purchase it or acquire it from a library. But what the searcher derived from the snippet was a historical fact. Author Goldberg's copyright does not extend to the facts communicated by his book. It protects only the author's manner of expression. *Hoehling v. Universal City Studios, Inc.*, 618 F.2d 972, 974 (2d Cir. 1980) ("A grant of copyright in a published work secures for its author a limited monopoly over the *expression* it contains.") (emphasis added). Google would be entitled, without infringement of Goldberg's copyright, to answer the student's query about the year Roosevelt was afflicted, taking the information from Goldberg's book. The fact that, in the case of the student's snippet search, the information came embedded in three lines of Goldberg's writing, which were superfluous to the searcher's needs, would not change the taking of an unprotected fact into a copyright infringement.

Even if the snippet reveals some authorial expression, because of the brevity of a single snippet and the cumbersome, disjointed, and incomplete nature of the aggregation of snippets made available through snippet view, we think it would be a rare case in which the searcher's interest *in the protected aspect* of the author's work would be satisfied by what is available from snippet view, and rarer still—because of the cumbersome, disjointed, and incomplete nature of the aggregation of snippets made available through snippet view—that snippet view could provide a significant substitute for the purchase of the author's book.

Accordingly, considering the four fair use factors in light of the goals of copyright, we conclude that Google's making of a complete digital copy of Plaintiffs' works for the purpose of

providing the public with its search and snippet view functions (at least as snippet view is presently designed) is a fair use and does not infringe Plaintiffs' copyrights in their books.

III. Derivative Rights in Search and Snippet View

Plaintiffs next contend that, under Section 106(2), they have a derivative right in the application of search and snippet view functions to their works, and that Google has usurped their exclusive market for such derivatives.

There is no merit to this argument. As explained above, Google does not infringe Plaintiffs' copyright in their works by making digital copies of them, where the copies are used to enable the public to get information about the works, such as whether, and how often they use specified words or terms (together with peripheral snippets of text, sufficient to show the context in which the word is used but too small to provide a meaningful substitute for the work's copyrighted expression). The copyright resulting from the Plaintiffs' authorship of their works does not include an exclusive right to furnish the kind of information about the works that Google's programs provide to the public. For substantially the same reasons, the copyright that protects Plaintiffs' works does not include an exclusive derivative right to supply such information through query of a digitized copy.

The extension of copyright protection beyond the copying of the work in its original form to cover also the copying of a derivative reflects a clear and logical policy choice. An author's right to control and profit from the dissemination of her work ought not to be evaded by conversion of the work into a different form. The author of a book written in English should be entitled to control also the dissemination of the same book translated into other languages, or a conversion of the

book into a film. The copyright of a composer of a symphony or song should cover also conversions of the piece into scores for different instrumentation, as well as into recordings of performances.

This policy is reflected in the statutory definition, which explains the scope of the "derivative" largely by examples—including "a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgement, [or] condensation"—before adding, "or any other form in which a work may be recast, transformed, or adapted." 17 U.S.C. § 101.[25] As noted above, this definition, while imprecise, strongly implies that derivative works over which the author of the original enjoys exclusive rights ordinarily are those that re-present the protected aspects of the original work, *i.e.*, its expressive content, converted into an altered form, such as the conversion of a novel into a film, the translation of a writing into a different language, the reproduction of a painting in the form of a poster or post card, recreation of a cartoon character in the form of a three-dimensional plush toy, adaptation of a musical composition for different instruments, or other similar conversions. If Plaintiffs' claim were based on Google's converting their books into a digitized form and making that digitized version accessible to the public, their claim would be strong. But as noted above, Google safeguards from public view the digitized copies it makes and allows access only to the extent of permitting the public to search for the very limited information accessible through the search function and snippet view. The program does not allow access in any substantial way to a book's expressive content. Nothing in the statutory definition of a derivative work, or of the

[25] The complete text is set forth at footnote 15, *supra*.

logic that underlies it, suggests that the author of an original work enjoys an exclusive derivative right to supply information about that work of the sort communicated by Google's search functions.

Plaintiffs seek to support their derivative claim by a showing that there exist, or would have existed, paid licensing markets in digitized works, such as those provided by the Copyright Clearance Center or the previous, revenue-generating version of the Google Partners Program. Plaintiffs also point to the proposed settlement agreement rejected by the district court in this case, according to which Google would have paid authors for its use of digitized copies of their works. The existence or potential existence of such paid licensing schemes does not support Plaintiffs' derivative argument. The access to the expressive content of the original that is or would have been provided by the paid licensing arrangements Plaintiffs cite is far more extensive than that which Google's search and snippet view functions provide. Those arrangements allow or would have allowed public users to read substantial portions of the book. Such access would most likely constitute copyright infringement if not licensed by the rights holders. Accordingly, such arrangements have no bearing on Google's present programs, which, in a non-infringing manner, allow the public to obtain limited data about the contents of the book, without allowing any substantial reading of its text.

Plaintiffs also seek to support their derivative claim by a showing that there is a current *unpaid* market in licenses for partial viewing of digitized books, such as the licenses that publishers currently grant to the Google Partners program and Amazon's Search Inside the Book program to display substantial portions of their books. Plaintiffs rely on *Infinity Broadcast*

Corporation v. Kirkwood, 150 F.3d 104 (2nd Cir. 1998) and *United States v. American Society of Composers, Authors and Publishers (ASCAP)*, 599 F. Supp. 2d 415 (S.D.N.Y. 2009) for the proposition that "a secondary use that replaces a comparable service licensed by the copyright holder, even without charge, may cause market harm." Pls.' Br. at 51. In the cases cited, however, the purpose of the challenged secondary uses was not the dissemination of information *about* the original works, which falls outside the protection of the copyright, but was rather the re-transmission, or re-dissemination, of their expressive content. Those precedents do not support the proposition Plaintiffs assert—namely that the availability of licenses for providing unprotected information about a copyrighted work, or supplying unprotected services related to it, gives the copyright holder the right to exclude others from providing such information or services.

While the telephone ringtones at issue in the *ASCAP* case Plaintiffs cite are superficially comparable to Google's snippets in that both consist of brief segments of the copyrighted work, in a more significant way they are fundamentally different. While it is true that Google's snippets display a fragment of expressive content, the fragments it displays result from the appearance of the term selected by the searcher in an otherwise arbitrarily selected snippet of text. Unlike the reading experience that the Google Partners program or the Amazon Search Inside the Book program provides, the snippet function does not provide searchers with any meaningful experience of the expressive content of the book. Its purpose is not to communicate copyrighted expression, but rather, by revealing to the searcher a tiny segment surrounding the searched term, to give some minimal contextual information to help the searcher learn whether the book's use of

that term will be of interest to her. The segments taken from copyrighted music as ringtones, in contrast, are selected precisely because they play the most famous, beloved passages of the particular piece—the expressive content that members of the public want to hear when their phone rings. The value of the ringtone to the purchaser is not that it provides information but that it provides a mini-performance of the most appealing segment of the author's expressive content. There is no reason to think the courts in the cited cases would have come to the same conclusion if the service being provided by the secondary user had been simply to identify to a subscriber in what key a selected composition was written, the year it was written, or the name of the composer. These cases, and the existence of unpaid licensing schemes for substantial viewing of digitized works, do not support Plaintiffs' derivative works argument.

IV. Plaintiffs' Exposure to Risks of Hacking of Google's Files

Plaintiffs argue that Google's storage of its digitized copies of Plaintiffs' books exposes them to the risk that hackers might gain access and make the books widely available, thus destroying the value of their copyrights. Unlike the Plaintiffs' argument just considered based on a supposed derivative right to supply information about their books, this claim has a reasonable theoretical basis. If, in the course of making an arguable fair use of a copyrighted work, a secondary user unreasonably exposed the rights holder to destruction of the value of the copyright resulting from the public's opportunity to employ the secondary use as a substitute for purchase of the original (even though this was not the intent of the secondary user), this might well furnish a substantial rebuttal to the secondary user's claim of fair use. For this reason, the *Arriba Soft* and *Perfect 10* courts, in upholding the secondary user's claim of fair use, observed that thumbnail

images, which transformatively provided an Internet pathway to the original images, were of sufficiently low resolution that they were not usable as effective substitutes for the originals. *Arriba Soft*, 336 F.3d 811 at 819; *Perfect 10*, 508 F.3d at 1165.

While Plaintiffs' claim is theoretically sound, it is not supported by the evidence. In *HathiTrust*, we faced substantially the same exposure-to-piracy argument. The record in *HathiTrust*, however, "document[ed] the extensive security measures [the secondary user] ha[d] undertaken to safeguard against the risk of a data breach," evidence which was unrebutted. 755 F.3d at 100. The *HathiTrust* court thus found "no basis . . . on which to conclude that a security breach is likely to occur, much less one that would result in the public release of the specific copyrighted works belonging to any of the plaintiffs in this case." *Id.* at 100-101 (citing *Clapper v. Amnesty Int'l USA*, 133 S.Ct. 1138, 1143 (2013) (finding that risk of future harm must be "certainly impending," rather than merely "conjectural" or "hypothetical," to constitute a cognizable injury-in-fact), and *Sony Corp.*, 464 U.S. at 453–454 (concluding that time-shifting using a Betamax is fair use because the copyright owners' "prediction that live television or movie audiences will decrease" was merely "speculative")).

Google has documented that Google Books' digital scans are stored on computers walled off from public Internet access and protected by the same impressive security measures used by Google to guard its own confidential information. As Google notes, Plaintiffs' own security expert praised these security systems, remarking that "Google is fortunate to have ample resources and top-notch technical talents" that enable it to protect its data. JA 1558, 1570. Nor have Plaintiffs identified any thefts from Google Books (or from the Google Library Project).

Plaintiffs seek to rebut this record by quoting from Google's July 2012 SEC filing, in which the company made legally required disclosure of its potential market risks.[26] Google's prudent acknowledgment that "security breaches could expose [it] to a risk of loss . . . due to the actions of outside parties, employee error, malfeasance, or otherwise," however, falls far short of rebutting Google's demonstration of the effective measures it takes to guard against piratical hacking. Google has made a sufficient showing of protection of its digitized copies of Plaintiffs' works to carry its burden on this aspect of its claim of fair use and thus to shift to Plaintiffs the burden of rebutting Google's showing. Plaintiffs' effort to do so falls far short.

---

[26] The filing includes the following disclosure:

> Our products and services involve the storage and transmission of users' and customers' proprietary information, and security breaches could expose us to a risk of loss of this information, litigation, and potential liability. Our security measures may be breached due to the actions of outside parties, employee error, malfeasance, or otherwise, and, as a result, an unauthorized party may obtain access to our data or our users' or customers' data. Additionally, outside parties may attempt to fraudulently induce employees, users, or customers to disclose sensitive information in order to gain access to our data or our users' or customers' data. Any such breach or unauthorized access could result in significant legal and financial exposure, damage to our reputation, and a loss of confidence in the security of our products and services that could potentially have an adverse effect on our business. Because the techniques used to obtain unauthorized access, disable or degrade service, or sabotage systems change frequently and often are not recognized until launched against a target, we may be unable to anticipate these techniques or to implement adequate preventative measures. If an actual or perceived breach of our security occurs, the market perception of the effectiveness of our security measures could be harmed and we could lose users and customers.

JA 562.

V. Google's Distribution of Digital Copies to Participant Libraries

Finally, Plaintiffs contend that Google's distribution to a participating library of a digital copy of Plaintiffs' books is not a fair use and exposes the Plaintiffs to risks of loss if the library uses its digital copy in an infringing manner, or if the library fails to maintain security over its digital copy with the consequence that the book may become freely available as a result of the incursions of hackers. The claim fails.

Although Plaintiffs describe the arrangement between Google and the libraries in more nefarious terms, those arrangements are essentially that each participant library has contracted with Google that Google will create for it a digital copy of each book the library submits to Google, so as to permit the library to use its digital copy in a non-infringing fair use manner. The libraries propose to use their digital copies to enable the very kinds of searches that we here hold to be fair uses in connection with Google's offer of such searches to the Internet public, and which we held in *HathiTrust* to be fair uses when offered by HathiTrust to its users. The contract between Google and each of the participating libraries commits the library to use its digital copy only in a manner consistent with the copyright law, and to take precautions to prevent dissemination of their digital copies to the public at large.

In these circumstances, Google's creation for each library of a digital copy of that library's already owned book in order to permit that library to make fair use through provision of digital searches is not an infringement. If the library had created its own digital copy to enable its provision of fair use digital searches, the making of the digital copy would not have been infringement. Nor does it become an infringement because, instead of making its own digital

copy, the library contracted with Google that Google would use its expertise and resources to make the digital conversion for the library's benefit.

We recognize the possibility that libraries may use the digital copies Google created for them in an infringing manner. If they do, such libraries may be liable to Plaintiffs for their infringement. It is also possible that, in such a suit, Plaintiffs might adduce evidence that Google was aware of or encouraged such infringing practices, in which case Google could be liable as a contributory infringer. But on the present record, the possibility that libraries may misuse their digital copies is sheer speculation. Nor is there any basis on the present record to hold Google liable as a contributory infringer based on the mere speculative possibility that libraries, in addition to, or instead of, using their digital copies of Plaintiffs' books in a non-infringing manner, may use them in an infringing manner.

We recognize the additional possibility that the libraries might incur liability by negligent mishandling of, and failure to protect, their digital copies, leaving them unreasonably vulnerable to hacking. That also, however, is nothing more than a speculative possibility. There is no basis in the record to impose liability on Google for having lawfully made a digital copy for a participating library so as to enable that library to make non-infringing use of its copy, merely because of the speculative possibility that the library may fail to guard sufficiently against the dangers of hacking, as it is contractually obligated to do. Plaintiffs have failed to establish any

basis for holding Google liable for its creation of a digital copy of a book submitted to it by a participating library so as to enable that library to make fair use of it.[27]

In sum, we conclude that: (1) Google's unauthorized digitizing of copyright-protected works, creation of a search functionality, and display of snippets from those works are non-infringing fair uses. The purpose of the copying is highly transformative, the public display of text is limited, and the revelations do not provide a significant market substitute for the protected aspects of the originals. Google's commercial nature and profit motivation do not justify denial of fair use. (2) Google's provision of digitized copies to the libraries that supplied the books, on the understanding that the libraries will use the copies in a manner consistent with the copyright law, also does not constitute infringement. Nor, on this record, is Google a contributory infringer.

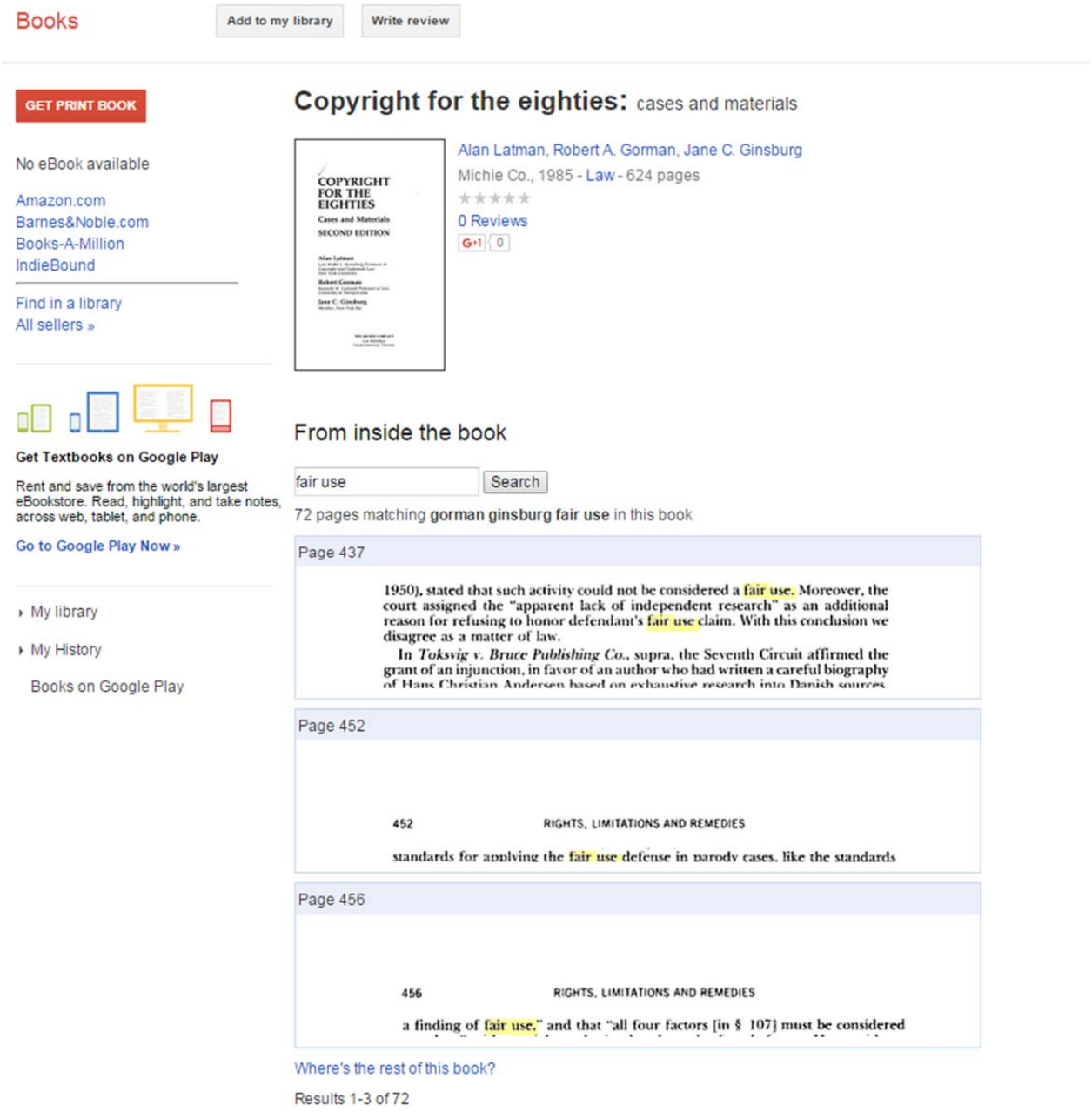
CONCLUSION

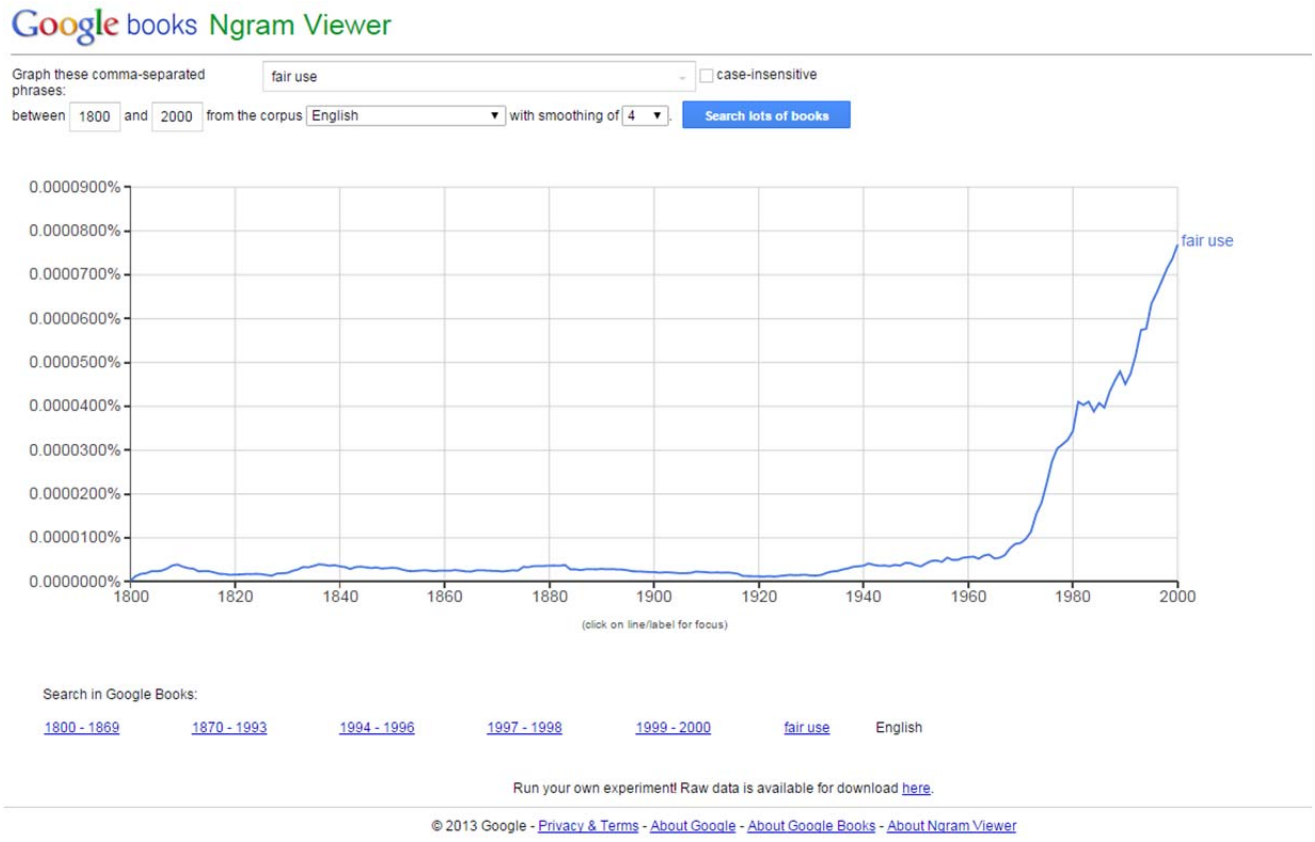
The judgment of the district court is **AFFIRMED**.

[27] We have considered Plaintiffs' other contentions not directly addressed in this opinion and find them without merit.

13-4829-cv
Authors Guild v. Google, Inc.

APPENDIX A



13-4829-cv
Authors Guild v. Google, Inc.

APPENDIX B