CORPORATE DISCLOSURE STATEMENT

The undersigned, counsel for Plaintiffs-Appellants, certifies that none of the Plaintiffs-Appellants have any parent corporation and no publicly-held corporation owns 10% or more of any of the Plaintiffs-Appellants' stock.

By: /s/ Edward H. Rosenthal
Edward H. Rosenthal
Jeremy S. Goldman
Anna Kadyshevich
Andrew D. Jacobs
FRANKFURT KURNIT KLEIN & SELZ, P.C.
488 Madison Avenue, 10th Floor
New York, New York 10022
(212) 980-0120

Paul M. Smith
JENNER & BLOCK LLP
1099 New York Avenue NW, Suite 900
Washington, D.C. 20001
(202) 639-6000

Attorneys for Plaintiffs-Appellants

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page(s)

Cases

Statutes

Other Authorities

PRELIMINARY STATEMENT

This case tells the story of how Google Inc. ("Google"), a commercial enterprise with virtually unlimited resources, enhanced its search engine, drove potential book purchasers away from online book retailers, increased its advertising revenue and stifled its competition by digitizing, distributing and monetizing millions of copyright-protected books without permission or payment.

When the mass digitization of books became technologically feasible, companies other than Google approached the opportunity with an appropriate respect for copyright. In 2003, Amazon asked publishers for permission to scan entire books and display up to twenty percent of those books in response to customer searches. (A56.)[1] After reaching agreements with over 190 publishers, including many of the largest in the United States, Amazon launched its "Search Inside the Book" program, with over 120,000 books and 33 million pages available for search and partial display. (A1292-93.) In order to use the service, Amazon required users to have an account with a credit card number attached to it so that the searchable books were only one click away from purchase. (A1293.) With this program, Amazon provided a retail environment where rightsholders could choose to make their books available for sale.

---

[1] As used herein, citations of the form "A___" refer to the Joint Appendix, "CA____" to the Joint Confidential Appendix, and "SPA___" to the Special Appendix.

Amazon's Search Inside the Book program threatened Google's dominance as *the* place people go to find information. (*See* A1294-98.) Indeed, just four months before Google's initial public offering in August 2004, Amazon launched its own search engine, A9.com, that included the materials licensed from rightsholders. But Google found a way to beat the competition by scanning copyright-protected books without permission from rightsholders. This way, it would avoid having to negotiate with publishers and authors—as Amazon did—about critical issues such as the extent of permitted text display, the security of the books and reasonable compensation. (*See* A1298-99.) In other words, Google would "neatly leapfrog" its competitors by ignoring copyright law and trampling the right of authors to control their works. (A1298.)

In December 2004, Google struck back, shocking the literary community by launching its unprecedented "Library Project," through which Google partnered with some of the world's largest libraries to gain free access to millions of copyright-protected books. Google emptied the shelves of libraries and delivered truckloads of printed books to scanning centers, where the books were converted into digital format. The resulting e-books were then copied onto Google's servers and ingested into its search engine. Google was savvy enough not to pay the libraries in cash for the incredible commercial advantage it gained from these

books. Instead, it provided the libraries with e-books worth millions of dollars as payment for their participation in the Library Project.

Google's ambitious Library Project was designed to lure potential book purchasers away from online retailers like Amazon—a website focused on selling books—and drive them to Google, a website focused on selling advertisements and keeping the resulting revenue. (CA440.) [redacted] Google's conduct also exposed the authors' works to new and serious security risks by converting their print books into digital format, storing the resulting e-books on multiple servers connected to the Internet and distributing millions of additional copies to the libraries, all without any promises to prevent loss of their property in the event of a data breach.

The District Court upheld Google's Library Project in its entirety, only a little more than two years after it had rejected a settlement that would have permitted Google's use of copyrighted works while providing rightsholders with revenue and security. At that time, the District Court had been broadly critical of Google's actions. *See The Authors Guild v. Google, Inc.*, 770 F. Supp. 2d 666, 669, 682-83 (S.D.N.Y. 2011) (observing that the settlement would reward Google for "engaging in wholesale copying of copyrighted works without permission").

The District Court concluded that fundamental changes in copyright law in the face of technological innovation are best left to Congress. *See id*. at 677-78 (citing *Eldred v. Ashcroft*, 537 U.S. 186, 212 (2003) ("it is generally for Congress, not the courts, to decide how best to pursue the Copyright Clause's objectives")). *Accord Sony Corp. of Am. v. Universal City Studios, Inc.,* 464 U.S. 417, 429 (1984).

Now the District Court has endorsed each and every aspect of Google's Library Project by applying an unprecedented, expansive and erroneous interpretation of the fair use doctrine. The District Court essentially ignored the inherently commercial nature of the Library Project and wrongly determined that the entire program is "transformative." In doing so, the District Court failed to separately evaluate whether each of Google's uses, including its reproduction, archival storage and distribution of the books in full, passes the transformative use test. The District Court further disregarded that, in order for Google's search function to operate, Google did not need to distribute the e-books to its Library Partners, display "snippets" of search results to its users or store multiple copies of the e-books on online servers. The District Court brushed aside the actual and potential harm to the literary market inflicted by the Library Project. And the District Court failed to heed one of the Supreme Court's key mandates in *Campbell v. Acuff-Rose Music, Inc*., 510 U.S. 569, 579 (1994), by failing to acknowledge that its ruling will open the door for others to copy, display and distribute digital

versions of copyrighted works for their own commercial advantages and subjecting those works to further security risks.

For the reasons set forth herein, Appellants request that this Court reverse the District Court's grant of summary judgment in favor of Google, grant Appellants' motion for partial summary judgment and hold Google liable for its copyright infringement. The case should be remanded to the District Court for the fashioning of an appropriate remedy. Authors wrote the books that fill the shelves of our bookstores and libraries. Appellants seek fair compensation for Google's commercial use of their books and for Google's distribution of their e-books to libraries. Google also must be required to remove its unauthorized digital library from servers connected to the Internet and otherwise to ensure that much of the world's literary heritage is not left susceptible to data breaches that would lead to widespread piracy and dissemination.

Google must not be permitted to build its financial empire on the backs of authors.

JURISDICTIONAL STATEMENT

The District Court had subject-matter jurisdiction over this action pursuant to 28 U.S.C. § 1338(a) because it arises under the copyright laws of the United States. (A177.) This Court has jurisdiction over this appeal pursuant to 28 U.S.C.

§ 1291. The District Court entered judgment on November 27, 2013 (SPA31-32), and an Amended Judgment entered on December 10, 2013 (SPA33-34).

The Notice of Appeal was timely filed on December 23, 2013. (A1639.)

STATEMENT OF ISSUES PRESENTED AND STANDARD OF REVIEW

I. Did the District Court err when it held that Google's unauthorized copying of millions of copyright-protected literary works as part of its Library Project is fair use under the Copyright Act?

Standard of Review: *De Novo*

II. Did the District Court err when it held that Google's unauthorized distribution of millions of copyright-protected works to its Library Partners is fair use under the Copyright Act?

Standard of Review: *De Novo*

III. Did the District Court err when it held that Google's unauthorized display of millions of copyright-protected literary works is fair use under the Copyright Act?

Standard of Review: *De Novo*

STATEMENT OF THE CASE

This is an appeal from the grant of summary judgment in favor of Google by the United States District Court for the Southern District of New York (Chin, J.) and the denial of the motion by the Appellants for partial summary judgment on

the issue of liability. (*See* SPA1-30); *Authors Guild, Inc. v. Google Inc.*, 954 F. Supp. 2d 282, 292 (S.D.N.Y. 2013). In the District Court, Appellants The Authors Guild, Inc., Betty Miles, Joseph Goulden and Jim Bouton (collectively, the "Authors") argued that Google's copying, distribution and display of millions of copyrighted books without permission of the copyright holders infringed their rights under the Copyright Act, 17 U.S.C. § 101 *et seq*. The District Court held that Google's entire Library Project is protected fair use and that Google owes no compensation to rightsholders for the commercial use of their books.

I. Google's Digital Books Program

A. Google's Licensed Partner Program

In October 2004, Google announced its original digital books program, calling it Google Print. (*See* CA3; CA40-41.) Significantly, Google Print included a program that displayed portions of books only with the permission of the publishers of each book (the "Partner Program"). (*See* A931 ¶¶ 22-24.)[2] The Partner Program allowed publishers and other rightsholders to permit Google to display their works in exchange for a split of ad revenue. Partners decided "how much of the book is browsable" on Google, "anywhere from a few sample pages

[2] As of early 2012, the Partner Program included approximately [redacted] books, by permission of approximately 45,000 rightsholders, with the number of partners continuing to grow. (CA92.) By early 2012, Google had phased out advertising adjacent to the display pages of books in the Partner Project. (CA1603.)

to the whole book." (A581.) In return, Google agreed to share with its partners a portion of the revenue it earned from ads shown next to pages of books searched in the Partner Program. (CA92.)

B. Google's Unlicensed Library Project

In December 2004, Google embarked on a new program to digitize millions of books without the permission of rightsholders. Google entered into agreements with several university libraries and the New York Public Library to "digitally scan books from their collections so that users worldwide can search them in Google." (A556.) In consideration for access to a library's print books, Google distributed copies of the resulting e-books to the contributing library. (A592; CA28.) Google referred to this endeavor as its "Library Project." (CA25-26; A583-85.) Since launching the Library Project, Google has entered into agreements with additional libraries, developed and patented scanning technology and copied over twenty million books. (A932 ¶¶ 27-29.)

1. Mass Digitization of Copyrighted Books

Google engaged in "bulk scanning," with libraries providing "carts of books" for Google to scan. (CA21, 43.) [redacted] Some libraries (*e.g.*, the New York Public Library, Harvard, Columbia, and Princeton) allowed Google to scan

only public domain works, while others (*e.g.*, the Universities of California, Michigan and Wisconsin) allowed Google to scan copyright-protected works as well. (*See* CA22; A593-684; CA116-308.)

In scanning upwards of four million books per year, Google did not independently assess the copyright status, commercial availability or content of any particular book. (CA63, 377.) Rather, Google copied every book its Library Partners provided, unless (a) Google determined that it already had copied or was scheduled to copy the book from another library, (b) the book was physically not fit to be copied, or (c) Google had received a specific request from a copyright owner not to scan the book. (CA316-17, 345, 377.)

Once a Library Partner's books were selected, they were shipped to Google's scanning centers to be digitized and indexed, a process that yielded numerous digital copies of each book. [redacted]

[redacted]

The scanning process resulted in an index that contains the complete text of all the books copied in the Library Project. The index is integrated into Google's all-purpose search index, so that regardless of whether a user submits a search via a general "Web" query or a "Books" query, Google will search the Library Project index and return results from Library Project books to the user. (CA48, 76.)

If a user clicks on a search result referring to a particular book, she will be taken to an "About the Book" page for that book. Since 2005, Google has displayed verbatim text from copyrighted books on these pages. (A699; CA384-85.) Google generally divides each page image into eighths, which it calls "snippets." (A706; CA383-84.) Once a user retrieves a book through her initial search, she can enter any other search terms she chooses, and the author's verbatim words will be displayed in three snippets for each search. Although Google has stated that any given search by a user "only" displays three snippets of each book (A699), a single user can view far more than three snippets from a

Library Project book by performing multiple searches using different terms, including terms suggested by Google. (*See* A445-509; CA28.) Even minor variations in search terms will yield different displays of text. (*Compare* A511-15, *with* A517-21; *see also* CA28.)[3] Google displays snippets from each book, except that it withholds display of 10% of the pages in each book and of one snippet per page. (A225 ¶ 10; CA383-84.) Thus, Google makes the vast majority of the text of these books—in all, 78% of each work—available for display to its users.[4]

Although snippets are shown in connection with user searches, they are *not* necessary to the indexing or other non-display functions upon which Google has sought to justify the Library Project. (*See* CA381-82; CA49.) Indeed, the Library Partners created a search engine for the e-books they received from Google that

---

[3] The Record includes examples of verbatim text displays for one book of each of the three named Authors. (*See* A982-1073 (222 unique snippets and ~11,611 words of text from Bouton's *Ball Four*); A1074-1171 (220 unique snippets and ~9,676 words of text from Goulden's *The Super-Lawyers*); A1172-1203 (61 unique snippets and ~1,760 words of text from Miles's *The Trouble with Thirteen*).)

[4] Google has disabled the display function for certain limited categories of works it has chosen to exclude, [redacted] In limiting the display of these books but not others, Google recognizes that in some instances showing even a "snippet" of a work may operate as a substitute for the original.

does not display any copyrighted book content. *See* HathiTrust Digital Library, http://www.hathitrust.org.

To date, Google has scanned over twenty million books through its Library Project, and it has made over four million in-copyright English language books available for snippet display. (A588; *see* CA25-26.) Google did not seek or obtain permission from the Authors or any other rightsholders before it made these uses, and it has not compensated them. (A701-03.)

2. Agreements with Library Partners

Pursuant to its agreements with its Library Partners, Google distributes digital copies of e-books to the providing library as payment for permission to scan. (CA28; CA504-05 ¶ 64.)[5] The providing library "owns" the digital copies distributed by Google. (CA505-06 ¶¶ 68-70.) As of March 26, 2012, Google had distributed 2.7 million of scanned books to the Library Partners. (A430.)

Google has little control over what the Library Partners do with their e-books; indeed, its library agreements presume the Library Partners' further use of the digitized works. (CA506 ¶ 71.)[6] For example, Google's agreement with the

[5] Google distributes copies to libraries through an online portal that allows Library Partners to request and download e-book versions of the books they provided to Google. (A396-7.) [redacted]

[6] The agreements make no effort to comply with Section 108 of the Copyright Act, which expressly governs permissible copying by libraries. Among other things,

University of Michigan specifically provides that Michigan's copies may be used for "inclusion in Michigan's search services" (A594); the university has since deposited those files—including files of works that are in-copyright—in its HathiTrust database index, which is accessible to anyone in the United States with an Internet connection (CA426, 429). Similarly, [redacted]

Section 108 requires that any copying or distribution be "made without any purpose of direct or indirect commercial advantage" (17 U.S.C. § 108(a)), allows copying for preservation purposes only where a work is unpublished (17 U.S.C. § 108 (b)), and allows copying for replacement of "damaged, deteriorating, lost or stolen" books only if new copies are not commercially available at a fair price (17 U.S.C. § 108(c)).

3. Google's Overriding Commercial Purposes

It is undisputed that the motivation behind Google's systematic digitization of millions of copyrighted works, [redacted]

Google operates the largest Internet search engine in the world, which it monetizes by displaying advertisements in response to search queries (CA46). Google reported over $36.5 billion in advertising revenues for 2011, up from $28 billion in 2010. (A563, 566-67.) The Library Project gives Google access to a vast universe of searchable data: the contents of millions of in-copyright books. Google's competitors, of course, do not have access to this material; on the contrary, they limited their own digitization projects to public domain or authorized works (*see* A56; A67). Thus, as Yahoo! explained in its objection to the proposed settlement in this case, the Library Project gives Google a distinct edge over competitors on its flagship product:

> With sole access to [its unauthorized scans], Google can better refine its search algorithm and gain a tremendous advantage against other search providers. . . . As a consequence of Google's singular position in the emerging book search marketplace . . . Google will gain a tremendous advantage in its core business area: Google Search.

(A81; *see also* A79; A941 ¶ 75-77.)

Indeed, Google launched the Library Project in direct response to Amazon's announcement that it was introducing a licensed full-text book search portal.

(A1298-99.) [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

4. Impact on the Authors' Books

Google's massive, unauthorized digitization campaign has acutely harmed the interests of the Authors and other rightsholders whose works have been copied. Most directly, Google has lured potential book buyers away from online bookstores and provided no compensation to rightsholders for Google's revenue-generating uses of their books (A702-03).

The Library Project also has impeded the emerging market for digital uses of books by search engines, libraries and others. Collective management organizations ("CMOs") typically pay rightsholders based on actual use of their works (A786-87), which can result in substantial revenue. For example, the Copyright Clearance Center ("CCC"), which licenses printed material of the same nature as that scanned by Google (A785), paid rightsholders more than $171 in 2011 (A791). Although collective licensing markets often have emerged

in response to new technologies (A785-786, 795), Google's indiscriminate and unauthorized uses "thwart the development of collective management systems for the digital use of book and book excerpts that authors and publishers would otherwise likely develop." (A796.) Significantly, there is strong evidence that a collective license approach would develop here, as evidenced both by the proposed settlement to this action (*see generally* A83-126) and by agreements reached abroad by Google, among others, regarding book digitization.[7]

Moreover, Google has subjected copyright-protected books to unacceptable security risk. Google acknowledges that its "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." (A562.) Worse, Google is not contractually obligated to, and does not in practice, monitor or control the security of the e-books it distributes to Library Partners. (*See* CA508 ¶ 108.) The risk of a potential security breach is exacerbated by this distribution, as the Library Partners' copies are stored on campus servers connected to the Internet—a prime target for hackers and activists.

---

[7] *See, e.g.*, Pls' Reply Br. Summ. J. at 16, ECF No. 1085 (citing Eric Pfanner, *Google Has Deal in France for Book-Scanning Project*, N.Y. Times, June 12, 2012, at B5); *see also* Alexis C. Madrigal, *Norway Decided to Digitize All the Norwegian Books*, The Atlantic, December 3, 2013, http://www.theatlantic.com/technology/archive/2013/12/norway-decided-to-digitize-all-the-norwegian-books/282008/.

(A802-03; *see also* Richard Pérez-Peña, *Campuses Face Rising Threat from Hackers*, N.Y. Times, July 17, 2013, at A1.; Wallace D. Loh, UMD Data Breach: Update #6 (Feb. 25, 2014), http://uhr.umd.edu/2014/02/umd-data-breach-update-6-02252014.) As the number of unlawful copies of an in-copyright book increases, so does the risk of further infringement and piracy of the work. (A801-02, 805-06.) In this way, even one security breach could devastate the literary market. (A806.)

II. Procedural History

A. The Lawsuit

This action was commenced on September 20, 2005. The original plaintiffs were The Authors Guild, which sued as an associational plaintiff on behalf of its members, as well as a number of individual authors, suing on their own behalf and on behalf of those similarly situated. Plaintiffs filed amended complaints on July 26, 2006 (ECF No. 36), October 31, 2008 (ECF No. 59), November 16, 2009 (ECF No. 782), and October 14, 2011 (ECF No. 985). The now-operative Fourth Amended Class Action Complaint alleges that Google's Library Project constitutes copyright infringement on a massive scale. (A176.)

The Authors Guild, the nation's largest membership organization of published authors, advocates for and supports the copyright and contractual interests of published writers. (A547.) It seeks equitable relief for authors affected

by Google's copyright infringement. The representative plaintiffs are the authors of published, in-copyright books that have been copied, distributed, and displayed by Google without the Authors' permission. Jim Bouton holds the copyright in *Ball Four* (A435-37); Betty Miles holds the copyright in *The Trouble With Thirteen* (A438-40); and Joseph Goulden holds the copyright in *The Superlawyers* (A441-43). The representative plaintiffs seek damages in addition to injunctive and declaratory relief.

B. Settlement History

Initial settlement discussions began in the fall of 2006. After more than two years of negotiations, the parties filed a proposed settlement agreement on October 28, 2008. The agreement was preliminarily approved by Judge John E. Sprizzo on November 17, 2008. (Order, ECF No. 64.) Notice of the proposed settlement triggered a number of objections, in response to which the parties entered into further negotiations. On November 13, 2009, the parties executed and filed a motion for approval of a proposed Amended Settlement Agreement ("ASA"). (*See* ECF No. 768.) Judge Chin entered an order preliminarily approving the ASA on November 19, 2009. (Order, ECF No. 772.)

Under the ASA, Google was to pay to plaintiff class members 63% of the revenues earned from specific, circumscribed uses, in the United States, of certain copyrighted works. (A85-87, 112-113.) Google also was to pay $34.5 million to

fund a Book Rights Registry, managed jointly by authors and publishers, that would locate rightsholders, maintain a database of their contact information, collect and pay revenues to the rightsholders for Google's use of works and otherwise protect and represent rightsholders' interests. (A85-87, 116-26.) Google further agreed to pay a minimum of $45 million to rightsholders in consideration for the release of Plaintiffs' claims for past copyright infringement. (A85.) Significantly, under the ASA, rightsholders were to retain control over the scanning, storage, display and sale of their works. In addition, because Google's rights were to be non-exclusive, rightsholders could license the same works to others, including Google's competitors. (A87-111.)

Notice of the ASA triggered another round of objections. The objections varied widely, raising copyright, antitrust, privacy and international law concerns. Judge Chin conducted a fairness hearing on February 18, 2010. On March 22, 2011, he rejected the ASA, holding that it "contemplates an arrangement that exceeds what the Court may permit under Rule 23" and thus failed to meet the "fair, reasonable, and adequate" standard. *Authors Guild*, 770 F. Supp. 2d at 667. Judge Chin emphasized that the issues raised by Google's mass digitization were "a matter more suited for Congress than this Court." *Id*. at 677.

C. Class Certification

On October 14, 2011, in the wake of the rejection of the ASA, Plaintiffs filed the Fourth Amended Class Action Complaint. (A175-90.) On December 12, 2011, the representative plaintiffs moved for class certification, and on December 22, 2011, Google moved to dismiss the Authors Guild's claims under Rule 12(b)(1) for lack of associational standing. On May 31, 2012, Judge Chin granted the representative plaintiffs' motion for class certification and denied Google's motion to dismiss. *Authors Guild v. Google, Inc.*, 282 F.R.D. 384 (S.D.N.Y. 2012).

On July 27, 2012, Plaintiffs moved for partial summary judgment on their copyright infringement claims, and Google cross-moved for summary judgment based on its fair use defense. (ECF Nos. 1031, 1044.) On September 17, 2012, before summary judgment briefing was completed, this Court stayed proceedings in the District Court pending Google's appeal of the class certification order. (ECF No. 1063.) On July 1, 2013, this Court, in a *per curiam* opinion, vacated the certification order, 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," and remanded the case "for consideration of the fair use issues." *Authors Guild, Inc. v. Google Inc.*, 721 F.3d 132, 134-35 (2d Cir. 2013). The parties then completed briefing on their summary judgment motions.

D. The Decision Below

In an opinion dated November 14, 2013, the District Court denied Plaintiffs' summary judgment motion and granted Google's motion, holding that "Google's use of the copyrighted works is 'fair use' under the copyright laws." (SPA16.) For the reasons set forth below, the District Court's decision must be reversed, summary judgment granted to the Authors and the case remanded to the District Court so that it can fashion a remedy that provides compensation to the Authors and security for their works.

SUMMARY OF ARGUMENT

The District Court erred in determining that Google's wholesale reproduction, storage, dissemination and display of millions of in-copyright works is fair under 17 U.S.C. Section 107. As shown herein, application of the four factors to each of Google's uses weighs against a finding of fair use.

As to first factor—the purpose and character of the use—the District Court did not accord any meaningful weight to Google's overriding commercial purpose, failed to consider each of Google's uses separately and erred in finding that such uses are "highly transformative." Next, in light of the vast number of highly creative and original works copied, distributed and displayed by Google, application of the second factor—the nature of the copyrighted work—also weighs against fair use. Google's indiscriminate copying, distribution and near-complete

display of copyrighted works plainly is far in excess of what it needed even to achieve the search and text mining functionality of the Library Project that the District Court found to be transformative. Accordingly, the third factor weighs heavily in the Authors' favor. Finally, as to the fourth factor, Google's uses inflict significant harm on the market and value of the Authors' works. Not only did Google fail to pay the Authors, but the Library Project undermines existing licensing markets and impedes the development of emerging ones. It also subjects the Authors' works to new risks of digital piracy and theft.

Because all four statutory factors weigh against fair use, Google's reproduction, distribution and display of e-books infringed the Authors' copyrights. Further, any change in the established balance between rightsholders and users should be made by Congress, not a private commercial entity like Google.

ARGUMENT

I.

THE DISTRICT COURT ERRED IN DETERMINING THAT THE LIBRARY PROJECT MAKES FAIR USE OF THE AUTHORS' COPYRIGHT-PROTECTED BOOKS

The District Court recognized that "Google has digitally reproduced millions of copyrighted books, including the individual plaintiffs' books," that "Google has made digital copies available for its Library Project partners to download," and that "Google has displayed snippets from the books to the public"—all "without license

or permission from the copyright owners." (SPA15-16.) Thus, it is undisputed that Google violated the Authors' exclusive rights of reproduction, distribution and display under Sections 106(1), (3) and (5) of the Copyright Act. The sole question on this appeal is whether the District Court erred in ruling that these acts are non-infringing because each and every part of Google's Library Project constitutes a fair use of the Authors' books.[8]

In the pages below, the Authors will show that the uses challenged here, including reproducing and storing the entirety of millions of copyrighted books, distributing those e-books to libraries as payment for access to their collections and displaying 78% of the e-books to the public—do not, and as a matter of sound copyright policy, *should* not constitute fair uses under the Copyright Act. While Google may have figured out a way to transform the creative output of authors into additional advertising dollars for Google's shareholders, Google's unauthorized commercial uses are not transformative within the meaning of copyright law.

Far from encouraging creative expression, Google has profited handsomely from the labors of the Authors while leaving them nothing for their efforts. Properly applied, all four factors militate strongly against a finding of fair use.

---

[8] Fair use is an affirmative defense as to which the alleged infringer has the burden of proof. *Campbell*, 510 U.S. at 590.

A. Purpose and Character of the Use

The first statutory factor, "the purpose and character of the use," 17 U.S.C. § 107(1), is "[t]he heart of the fair use inquiry." *On Davis v. The Gap, Inc.*, 246 F.3d 152, 174 (2d Cir. 2001). In considering this factor, a court examines "whether the new work merely 'supersede[s] the objects' of the original creation . . . or 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.'" *Campbell*, 510 U.S. at 579 (citations omitted). By express statutory instruction, the Court must also consider "whether the use is of a commercial nature" or whether it is "for nonprofit educational purposes." 17 U.S.C. §107.

1. Google's Uses Are Highly Commercial

The District Court failed to give any meaningful weight to the undisputed fact that the nature of Google's Library Project is highly commercial. At most, the court gave lip service to the statutory language, merely mentioning "[t]he fact that a use is commercial 'tends to weigh against a finding of fair use,'" (SPA21 (quoting *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 562 (1985))), and then essentially disregarding it.

The District Court, however, trivialized Google's overriding commercial purpose. While acknowledging that Google "benefit[s] commercially in the sense

that users are drawn to the Google websites by the ability to search Google books," it stated that Google "does not sell the scans it has made of books[,] . . . does not sell the snippets that it displays[,] and . . . does not run ads on the About the Book pages that contains snippets." (SPA21-22.) It found support for its position in cases that could hardly be more different than the case before this Court—cases involving creative secondary uses of a single or small number of copyrighted works. In its ruling, the District Court completely ignored evidence demonstrating that (a) Google's embarked on its Library Project for the sole purpose of gaining a competitive advantage and increasing revenues, and (b) the scope of Google's Library Project dwarfs that of any prior fair use case.

Google's brazen mass digitization project was designed and has helped Google to dominate the search engine market and reap vast corporate profits. [redacted]

As Yahoo! explained in its objection to the ASA, "the more data available . . . the better the search engine" (A81); as a result, the Library Project has given Google "a tremendous advantage in its core business area: Google Search." (*Id.*; *see also* A1298-99 (Google began digitizing copyright-protected books in direct response to Amazon's announcement that it was launching a licensed full-text book search portal).)

The District Court makes much of the fact that Google "does not run ads on the About the Book pages that contain snippets." (SPA21-22.) This reflects a fundamental misunderstanding of Google's business model. The main source of Google's advertising profits does not come from the "About the Book" pages, but rather the ad-laden search results pages that link to them. When Google's millions of U.S. users come to the site every month to locate relevant information, the results pages display content from, and links to, *all* of the content in Google's massive database, including content ingested from the Library Project. In many instances, this content runs alongside paid advertisements tied to the search results. For example, a search for "Steve Hovley" yielded not only an excerpt from Plaintiff-Appellant Jim Bouton's *Ball Four*, but also paid advertisements for three public records search services and eBay. (A242.) In this way, Google derives direct commercial benefit from the Authors' words, phrases, and creative expression.

Google also uses the copyrighted material for myriad commercial purposes beyond advertising, all designed to enhance the functionality of Google's services and, as a result, increase its dominance in the information technology market. As one commentator explained in connection with the proposed settlement:

> It seems likely that the 'nondisplay uses' of Google's scanned corpus of text will end up being far more important than anything else in the agreement. Imagine the kinds of things that data mining all the world's books might let Google's engineers build: automated translation, optical character recognition, voice recognition algorithms. And those are just the things we can think of today. Under the agreement, Google has unrestricted, royalty-free access to this corpus.

Fred von Lohmann, Google Book Search Settlement: A Reader's Guide, Electronic Frontier Foundation Deep Links Blog (Oct. 31, 2008), http://www.eff.org/deeplinks/2008/10/google-books-settlement-readers-guide.

Thus, while its competitors went through the "painstaking" and "costly" process of obtaining permission before scanning copyrighted books, "'Google by comparison took a shortcut by copying anything and everything regardless of copyright status.'" *Authors Guild*, 770 F. Supp. 2d at 679 (quoting Hr'g Tr. 43 (Thomas Rubin, counsel for Microsoft)). "As one objector put it: 'Google pursued its copyright project in calculated disregard of authors' rights. Its business plan was: So, sue me.'" *Id*. (quoting Obj. of Robert M. Kunstadt to Proposed Settlement at 3, ECF No. 74).

Because this case involves the wholesale digitization and use of *millions* of works for commercial purposes, it is vastly different from the many fair use cases that challenged an artist's or author's unauthorized use of one or a small handful of copyrighted works for the purpose of creating a new work of creative expression. *See, e.g.*, *Campbell*, 510 U.S. at 571-72 (rap group's use of iconic guitar riff and lyrical refrain to create a parody of the original); *Cariou v. Prince*, 714 F.3d 694, 699-702 (2d Cir. 2013) (appropriation artist's use of photographs to create new artwork); *Blanch v. Koons*, 467 F.3d 244, 246 (2d Cir. 2006) (visual artist's use of single photograph in creating a new collage painting); *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 607 & n.1 (2d Cir. 2006) (use of seven *Grateful Dead* concert posters in book about the band).

In *Am. Geophysical Union v. Texaco Inc.*, 802 F. Supp. 1 (S.D.N.Y. 1992), *aff'd*, 60 F.3d 913, 923 (2d Cir. 1995), then-District Judge Pierre Leval observed an analytic distinction between "productive" copying cases (*i.e.*, the "classic" fair use cases like those described above that involve the limited use of a pre-existing work to create something truly new) and "nonproductive" copying cases (*i.e.*, less common cases that involve mechanical copying on a larger scale). Judge Leval noted that, in the wake of the Supreme Court's ruling in *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417 (1984), courts had held that certain

nonproductive copying could be fair use, *id*. at 12-13, but only where such uses were noncommercial:

> What has emerged . . . seems to be a two-track pattern of interpretation of the first factor: Secondary users have succeeded in winning the first factor by reason of either (1) transformative (or productive) nonsuperseding use of the original, or (2) noncommercial use, generally for a socially beneficial or widely accepted purpose.

Id. at 12. For classic "productive" fair use cases, the transformative use test has dominated the first-factor inquiry. *See id*. By contrast, where the challenged use adds nothing new to the original by "altering the first with new expression, meaning, or message," *Campbell*, 510 U.S. at 579, the commerciality of the use is paramount. *See Texaco*, 802 F. Supp at 13.

Here, the District Court erred by employing a fair use analysis designed and best suited for cases involving use of a single or small number of copyrighted works. Thus, the District Court's reliance on cases like *Blanch* and *Bill Graham Archives*, in which the challenged use was deemed fair despite having a commercial aspect (*i.e*., selling artwork or books), is misplaced. (SPA20-21.) In those cases, an artist or author took a small number of pre-existing works for the purpose of creating a new work of expression. By contrast, Google used millions of pre-existing works to build its search engine and drive advertising revenue. Google's commercial enterprise stands in stark contrast to the lone visual artist creating a collage in his studio out of images scanned from glossy magazine

photographs, *see Blanch*, 467 F.3d at 247, or a Grateful Dead enthusiast collecting concert posters to include in a biographical anthology about the band, *see Bill Graham*, 448 F.3d at 607.

When considering the commercial aspect of Google's enterprise, a far more analogous case is *Princeton Univ. Press v. Michigan Document Servs., Inc.*, in which the Sixth Circuit held that the duplication of copyrighted materials by a copy shop was "for sale by a for-profit corporation that has decided to maximize its profits—and give itself a competitive edge over other copy shops—by declining to pay the royalties requested by the holders of the copyrights." 99 F.3d 1381, 1386 (6th Cir. 1996). *See also United States v. ASCAP*, 599 F. Supp. 2d 415, 429 (S.D.N.Y. 2009) ("Because the [free ringtone] previews constitute a form of advertisement for applicant's ringtones and ringback tones . . . , applicant's use of ASCAP music to increase revenues from sales of ringtones and ringback tones is commercial."); *UMG Recordings, Inc. v. MP3.com, Inc.*, 92 F. Supp. 2d 349, 351 (S.D.N.Y. 2000) (commercial use where defendant offered free subscriptions to its service and did not sell the songs it had illegally copied in order "to attract a sufficiently large subscription base to draw advertising and otherwise make a profit").

Contrary to the District Court's holdings, the commerciality of Google's uses weighs very strongly against a finding of fair use.

2. Google's Uses Are Not Transformative

The reproduction, distribution and display of millions of in-copyright works do not constitute transformative uses.[9] In *Campbell*, the Supreme Court stated that a use is transformative when it "adds something new, with a further purpose or different character, altering the first with *new expression, meaning, or message*." 510 U.S. at 579 (1994) (emphasis added). The requirement that the use have "a further purpose or different character" does not stand alone, but instead modifies the ultimate requirement that the secondary use "adds something new." *Id.* As a result, merely articulating a new "purpose" for a use of a copyrighted work, without changing or adding anything new, is not enough to render the work transformative. *See id.*; *Cariou*, 714 F.3d at 706 (while a work need not "comment on the original or its author in order to be considered transformative," it still must "alter[] the original with new expression, meaning or message"). Indeed, if a

---

[9] These uses must be analyzed separately. Section 107 of Copyright Act requires an examination of "whether *the use* made of a work in any particular case is a fair use." 17 U.S.C. § 107 (emphasis added). It does not ask whether all of an infringer's various uses, taken in the aggregate, are fair. *Cf. Cariou*, 714 F.3d at 711-12 (holding "all except five" of alleged infringer's artworks to be fair use and remanding remaining five for further analysis); *Associated Press v. Meltwater U.S. Holdings, Inc.*, 931 F. Supp. 2d 537, 557 (S.D.N.Y. Mar. 21, 2013) ("The fact that [defendant] also offers a number of analysis tools does not render its copying and redistribution of article excerpts transformative."); *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1169-73 (9th Cir. 2007) (upholding as fair use one of Google's uses (thumbnail displays of images), but determining that Google could be liable for another use (in-line linking to full-sized infringing images)).

"further purpose" alone were sufficient, a secondary user could always claim transformativeness merely by articulating some different purpose when infringing on a copyright. "Such a strategy empties the term of meaning—for the 'transformative' moniker to guide, rather than follow, the fair use analysis, it must amount to more than a conclusory label." 4 Melville B. Nimmer & David Nimmer, *Nimmer On Copyright* § 13.05[A][1][b] (2011).

In its analysis, the District Court seized on certain features of the Library Project, labeled those uses "transformative," and then applied that label to the entirety of the Library Project. In doing so, the Court transformed "transformative" into exactly the kind of conclusory label Professor Nimmer warned against.

(a) Reproducing and Storing Exact Digital Copies Are Not Transformative Uses

Google has scanned over four million in-copyright print books into e-books and has permanently stored multiple copies of those e-books on its servers. This use—mechanical conversion—is not transformative. Digitization, like photocopying, "merely transforms *the material object* embodying the intangible article that is the copyrighted original work." *Texaco*, 60 F.3d at 923 (emphasis added); *see Princeton Univ. Press*, 99 F.3d at 1389 ("mechanical 'transformation'" of photocopying "bears little resemblance to the creative metamorphosis accomplished by the parodists in the *Campbell* case"). Converting print books into

e-books adds no "new information, new aesthetics, [or] new insights and understandings," to the copyrighted books. Pierre N. Leval, *Toward A Fair Use Standard*, 103 Harv. L. Rev. 1105, 1111 (1990). *See Infinity Broad. Corp. v. Kirkwood*, 150 F.3d 104, 108 (2d Cir. 1998) (retransmission of radio broadcast, which merely repackages the original, is not transformative because it adds "neither new expression, new meaning nor new message") (quotation omitted). While mass digitization operations "may be innovative, they are not transformative." *UMG Recordings, Inc. v. MP3.Com, Inc.*, 92 F. Supp. 2d at 351 (claim that the conversion of CDs into computer files is transformative "is simply another way of saying that the unauthorized copies are being retransmitted in another medium—an insufficient basis for any legitimate claim of transformation").

The District Court sidestepped Google's reproduction and creation of millions of e-books, instead improperly focusing its analysis of transformativeness on the Library Project's independent search index and text mining functions. In determining that the Library Project "transforms expressive text into a comprehensive word index" and "has transformed book text into data for purposes of substantive research," it concluded that "Google's use of the copyrighted works is highly transformative." (SPA19-20.) However, to the extent that either of those purposes fits within the rubric of fair use (and it is the Authors' contention that

they do not), Google should not be permitted to make uses unnecessary to those purposes. *See Campbell*, 510 U.S. at 589 (remanding to determine whether defendant's repetition of the copyrighted song's bass riff was excessive copying in light of the song's parodic purpose). Because Google's permanent storage of multiple digital copies of every digitized book on its servers is entirely unnecessary for purposes of the Library Project's search index and text mining functions (*see* CA70, 352-53, 381-82), those purposes cannot render Google's reproduction transformative.

The Ninth Circuit cases cited by the District Court, *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007), and *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003), which address the legality of copying web pages for the purpose of creating a search index, are easily distinguishable. Unlike Google's perpetual storage of high resolution image files and text files of every book it scans, the images copied by the search engines in *Kelly* and *Perfect 10* were not permanently stored. *See Kelly*, 336 F.3d at 815 (after copying full size images onto its server for the purpose of creating "thumbnails," the search engine *deletes the original copy* from its server) (emphasis added); *Perfect 10, Inc.*, 508 F.3d at 1156-57 (noting that "Google does not store the images" and that "Google's cache contains the text of the webpage as it appeared at the time Google indexed the page, *but does not store images from the webpage*") (emphasis added). Here,

Google not only copies books for the purpose of ingesting the content into its search index, but it also permanently stores complete copies of the books for future exploitation and distribution to the libraries. Even if the Ninth Circuit considered the mechanical indexing of the words in a book to be "transformative," it does not follow that Google's reproduction, storage and distribution of unaltered copies of the books also is transformative. *See Texaco*, 60 F.3d at 924 ("In this case, the predominant archival purpose of the copying tips the first factor against the copier, despite the benefit of a more usable format.").

(b) Disseminating Millions of Copyrighted Books to Google's Library Partners is a Not Transformative Use

Google has distributed a vast number of copyright-protected e-books to its Library Partners. As of March 2012, the Library Partners had received over 2.7 million e-books. (A430.) Google is obligated to provide these e-books as in-kind payment for the Library Partners' allowing Google to digitize and commercially exploit their print collections. (CA28; CA504-05 ¶ 64.) Like Google's other challenged uses, this distribution-as-payment does not "alter the original[s] with 'new expression, meaning or message'" and thus is plainly non-transformative. *Cariou*, 714 F. at 706.

To the extent the District Court addressed this issue in isolation, it did so only with reference to the purposes of the Library Partners to which Google has distributed its digital copies. It held that "Google's actions constitute fair use [with

regard to distribution] as well," relying entirely on the determination that "the libraries . . . use these digital copies in transformative ways." (SPA26-27.) But the purposes of Google's end users, as opposed to those of Google, are irrelevant to a fair use analysis. *See Infinity*, 150 F.3d at 108 ("[I]t is [defendant's] own retransmission of the broadcasts, not the acts of his end-users, that is at issue here.").

Like the radio re-transmitter in *Infinity*, who sold "access to unaltered radio broadcasts," *id.*, Google has sold to its Library Partners access to unaltered copyrighted books. Moreover, pursuant to their agreements with Google, the Library Partners are not restricted to the uses that Judge Baer found transformative in *Authors Guild, Inc. v. HathiTrust*, 902 F. Supp. 2d 445 (S.D.N.Y. 2012), and on which Judge Chin relied in his analysis. Indeed, the agreement between Stanford and Google contemplates possible commercial uses of unauthorized e-books. (CA133-34.) Accordingly, there is a "total absence of transformativeness in [defendant's] acts." *Infinity*, 150 F.3d at 109.[10]

[10] Even if the Court determines that the Library Partners' uses are relevant when considering the fairness of Google's uses, the Authors disagree that the libraries' conduct, which is in direct contravention of the restrictions in Section 108 of the Copyright Act, is non-infringing. The Authors Guild and other authors and rights organizations have appealed Judge Baer's dismissal of their copyright claims against certain of Google's Library Partners. That appeal now is *sub judice* before this Court. *See Authors Guild, Inc. v. HathiTrust*, No. 12-4547.

(c) Displaying Large Portions of Millions of Copyrighted Books is Not Transformative

Google's display of 78% of the verbatim text of millions of in-copyright books is non-transformative because it does not "alter the original[s] with 'new expression, meaning or message.'" *Cariou*, 714 F. 3d at 706 (quoting *Campbell* at 589). Indeed, this Court has declined to find the use of verbatim excerpts transformative even where the use was less extensive and more creative than Google's unauthorized display. *See Castle Rock Entm't v. Carol Publ'g Grp.*, 150 F.3d 132, 141-143 (2d Cir. 1998) (use of forty-one verbatim quotations in trivia book that "involve[d] some creative expression" held non-transformative); *Twin Peaks Prods., Inc. v. Publications Int'l, Ltd.*, 996 F.2d 1366, 1372-77 (2d. Cir. 1993) (chapter of trivia questions featuring one page of verbatim quotations from television series held non-transformative).

The District Court erred in accepting Google's argument that helping readers "determine whether [a particular book] may be of interest" is a "different purpose for the book." (SPA20.) While the unauthorized display of snippets may incidentally inform readers whether a particular book is of interest, allowing users to view verbatim copyrighted material is a non-transformative and infringing substitute. *See, e.g., HarperCollins Publishers L.L.C. v. Gawker Media LLC*, 721 F. Supp. 2d 303, 306 (S.D.N.Y. 2010) (blog that displayed excerpts from forthcoming book had "not used the copyrighted material to help create something

new" but "merely copied the material in order to attract viewers"); *ASCAP*, 599 F. Supp. 2d at 424 ("Applicant's use of [ringtone] previews is not transformative. It is undisputed that the music segments used in applicant's previews are exact copies of ASCAP music."); *see also Texaco,* 60 F.3d at 923 ("[A]n untransformed copy is likely to be used simply for the same intrinsic purpose as the original"). According to the District Court's logic, any otherwise non-transformative use would be deemed fair so long as it also attracts readers to the original.

Similarly, the District Court's reliance on the *Bill Graham* case in this Circuit and *Perfect 10* and *Kelly* from the Ninth Circuit is misplaced. *Bill Graham* involved a 480-page book about the Grateful Dead. *Id*. at 607. The book included 2,000 images presented in chronological order, combining those images with text and graphic art to convey the history of the band. *Id*. At issue were seven of those images, photographs of concert posters greatly reduced in size for the book format and interspersed with narrative text and graphic art about those concerts. *Id*. This Court held the use transformative because "enhancing the biographical information in *Illustrated Trip* [was] a purpose separate and distinct from the original artistic and promotional purpose for which the images were created." *Id*. at 610. Here, by contrast, Google has displayed copyrighted language verbatim without incorporating it into any new work. Google did not, for example, publish an original biography about Jim Bouton's major league baseball career, using the

cover page and brief excerpts from *Ball Four* as part of that biography to illustrate that career.

Perfect 10 and *Kelly* each held that a search engine's display of thumbnail images was transformative. In those cases, however, the copyright holders voluntarily placed their content on publicly-accessible webpages which, by default, are copied by "web crawlers" and ingested into their dispatchers' search engines. Indeed, webpages are there to be located and accessed. Here, the Authors did not upload the content of their books onto the Web to be copied, indexed, stored and disseminated by search engines.

The Third Circuit case, *Video Pipeline, Inc. v. Buena Vista Home Entm't, Inc.*, 342 F.3d 191 (3d Cir. 2003), is far more analogous to this case. There, the defendant compiled movie trailers and clips and copied them into a database without authorization. *Id.* In rejecting the defendant's reliance on *Kelly* to argue that its database of video clips was fair use, the Third Circuit held:

> Video Pipeline's database does not, however, serve the same function as did Arriba Soft's search engine. As used with retailers' websites, VideoPipeline.net does not improve access to authorized previews located on other websites. Rather, it indexes and displays unauthorized copies of copyrighted works. VideoDetective.com does permit viewers to link to legitimate retailers' web sites, but a link to a legitimate seller of authorized copies does not here, if it ever would, make prima facie infringement a fair use.

Id. at 199; *see also id*. at 200 (Video Pipeline's clip previews lack any significant transformative quality and the commercial nature of the clip previews weighs more strongly against their use).

Finally, even if certain of Google's uses of copyrighted books are found to be transformative, the inquiry must not end there. As the *Campbell* court made clear, the extent to which a new work transforms the copyrighted work is simply a factor that is balanced against the other fair use factors. 510 U.S. at 577-78. Here, Google gained an enormous commercial and competitive advantage from its use of others' copyrighted works and, as described above and reiterated in the analysis of factor four below, stopped the emergence of a licensed market (already begun by Amazon) in its tracks. Google could have paid for the right to include copyrighted works in its search engine, as it agreed to do in the ASA, or worked to establish a collective licensing scheme of the sorts now being adopted in other countries. Google should not be permitted to reap enormous financial rewards on the backs of authors even if the Court finds that making books searchable is transformative.

B. Nature of the Copyrighted Work

Just as the District Court essentially wrote out of the statute the explicit requirement that the commercial nature of the use be considered as part of the first factor, it essentially eliminated factor two. The second statutory factor "calls for recognition that some works are closer to the core of intended copyright protection

than others, with the consequence that fair use is more difficult to establish when the former works are copied." *Campbell*, 510 U.S. at 586. While the District Court found that the "vast majority" of the books in the Library Project are non-fiction and published (SPA22-23), the works digitized, distributed and displayed as part of the Library Project include every conceivable type of book, including fiction and non-fiction, obscure academic works and bestsellers. In making its selections, Google gave no consideration to the nature of the books copied. (A932-33 ¶¶ 32-34.) Under the unusual and unprecedented circumstances of this case, where Google indiscriminately scanned millions of books in assembly-line fashion, Google cannot meet its burden of showing that the nature of the copyrighted work favors a finding of fair use.

The proposition that the second factor can be found to favor this massive *prima facie* infringement is troubling. As this Court recognized in the context of a single author's copyrighted works, "there is no easy distinction between works that are factual on the one hand, and 'creative' or expressive on the other, because creation of a nonfiction work, even a compilation of pure fact, entails originality." *New Era Publ'ns Int'l., ApS v. Carol Publ'g Grp.*, 904 F.2d 152, 158 (2d Cir. 1990) (internal citation and quotation marks omitted). Had Google scanned the entirety of a single work of fiction, the second factor presumably would have weighed in the favor of the rightsholder. The answer cannot be different where

Google scanned hundreds of thousands of works of fiction, along with a vast number of non-fiction works.

C. Amount and Substantiality of the Portion Used

The District Court's disregard of the second fair use factor pales in comparison with its treatment of the third factor, which considers "the amount and substantiality of the portion used in relation to the copyrighted work as a whole," or whether "the quantity and value of the materials used[] are reasonable in relation to the purpose of the copying." *Campbell*, 510 U.S. at 586 (internal quotation marks and citations omitted). This factor "recognizes that the more of a copyrighted work that is taken, the less likely the use is to be fair, and that even a less substantial taking may be unfair if it captures the essence of the copyrighted work." *Infinity*, 150 F.3d at 109. "Though not an absolute rule, 'generally, it may not constitute a fair use if the entire work is reproduced.'" *Id*. (quoting Nimmer, § 13.05[A][3] at 13-178).

Here, the third fair use factor takes on particular significance because truckloads of books were digitized, copied and disseminated in their entirety. 78% of each of these books was made available for public display. But while the District Court correctly found that "Google scans the full text of books—the entire books," it concluded that "the third factor weighs *slightly* against a finding of fair use." (SPA23-24) (emphasis added). It held that "full-work reproduction is

critical to the functioning of Google Books" and that because "Google limits the amount of text it displays in response to a search," Google's unbounded uses were mitigated. (SPA23.) Neither of these considerations stand to reason: the former is simply not true, and the latter ignores quantitative and the qualitative substantiality of snippets.

1. Google's Uses Far Exceed That Necessary for the Search and Text Mining Functions

The District Court's reference to "the functioning of Google Books" is vague; given that its analysis of Google's purpose centers entirely on the search and text mining functions, it may be assumed that it was referring to those functions in finding that "full-work reproduction is critical" here. (SPA23.) However, even accepting, *arguendo*, that it might be fair use to digitize books as an intermediate step in the creation of a searchable index and text mining database, the facts show that Google has gone much further than necessary to accomplish these purposes. *Cf. Campbell*, 510 U.S. at 589 (remanding case to determine whether the repetition of original song's iconic bass riff was excessive).

Google made and has kept multiple copies of the image and text files, but it has not demonstrated any search-related purpose served by the retention of such files after they are ingested into the search index. It made digital copies of the books available to its Library Partners (A396-97), not in furtherance of its search project, but rather as in-kind payment for access to the libraries' massive

collections of print books (CA28; CA504-05 ¶ 64). As part of the search function, Google displays significant portions of the books, even though its Library Partners concluded that no such display was required to fulfill the index and search purposes. *HathiTrust*, 902 F. Supp. 2d at 448. Far from being "critical" to the search index and text mining functions, the reproduction, distribution and display of e-books are completely unnecessary to the functioning of the search engine.

In short, the District Court minimized Google's taking of copyrighted works by focusing on one aspect of Google's program—search—in isolation, and ignoring the broader picture. As explained above, separate uses must be analyzed separately, and courts have refused to allow an activity that may qualify as fair use to excuse infringing activities that do not. The District Court erred in allowing Google's search function to excuse all its other uses of the Authors' copyrighted materials.

2. Snippets May Contain the "Heart" of the Book

The District Court's search-centric analysis also glosses over the great extent to which Google displays the Authors' works, minimizing the significance of this taking by relying on the bare fact that "Google limits the amount of text it displays in response to a search." (SPA23.) Putting aside that Google collectively displays 78% of each book, even single snippets are highly expressive and indeed may take "what [is] essentially the heart of the book." *Harper & Row*, 471 U.S. at 564-65

(internal quotation marks omitted); *see id*. at 565 (that quoted passages "qualitatively embodied [the author's] distinctive expression" weighs against finding of fair use); *see also HarperCollins Publishers L.L.C*, 721 F. Supp. 2d at 305-06 ("images of portions of 12 pages" of memoir "amounts to a substantial portion of the Book").

By disabling snippet view for recipes, cookbooks, and other reference works, Google has implicitly acknowledged the threat that public display of even short excerpts may pose to the "heart" of a book. But why should Google be allowed to make the decision as to which works fall into which category and which categories present a risk to Authors' rights? Google's own examples of allegedly lawful display from *Ball Four*, Jim Bouton's memoir about Steve Hovley, demonstrate the important difference between a mere index and the display of expressive text from a copyrighted work:

> Steve Hovley has been called up. Old Tennis Ball Head hasn't had a haircut since he left. Which means Joe Schultz had four comments to him the very first day. "Where's your barber?" "Don't you need a haircut?"
>
> * * *
>
> Moments later, perhaps feeling bad about his comment, or perhaps wishing to stick the needle in further, Pattin approaches Hovley and says, "I don't care you long you wear your hair, Hovley. You can wear it down to your ass as far as I'm concerned."

(A246.)

If the Library Project were merely to enable a user to search for books that reference "Steve Hovley," or view the number of times that the name "Steve Hovley" appears in a book, that would be one thing. But permitting users to read excerpts that bring to bear all of the author's originality, subjective insights, creative spirit and unique voice—is quite another. *See Heim v. Universal Pictures Co.*, 154 F.2d 480, 488 (2d Cir. 1946) (a single phrase may be sufficiently "idiosyncratic in its treatment" so as to receive protection). These, of course, are but two examples. Google users are able to view other displays of expression from millions of in-copyright books *ad infinitum*. Google's appropriation and display of expression has occurred on a massive, unprecedented scale that goes far beyond what is meaningfully connected to search and indexing functions.

D. The Effect of the Use Upon the Market for or Value of the Work

The fourth fair use factor, "the effect of the use upon the potential market for or value of the copyrighted work," 17 U.S.C. § 107(4),

> requires courts to consider not only the extent of market harm caused by the particular actions of the alleged infringed, but also "whether unrestricted and wide-spread conduct of the sort engaged in by defendant . . . would result in a substantially adverse impact on the potential market" for the original.

Campbell, 410 U.S. at 590 (quoting Nimmer, § 13.05[A][4]). The copyright holder need not show either "actual present harm" or "with certainty that future harm will result," but rather "a showing by a preponderance of the evidence that

some meaningful likelihood of future harm exists." *Sony*, 464 U.S. at 451 (emphasis in original). Courts must consider not only any damage caused to the original work, but also any harm to the market for derivative uses. *Harper & Row*, 471 U.S. at 568. As with the other fair use factors, Google bears the burden of showing that its use does not harm the Authors' interests. *Infinity*, 150 F.3d at 110.

Speculating that "Google Books enhances the sales of books to the benefit of copyright holders," the District Court concluded that "the fourth factor weighs strongly in favor of a finding of fair use." (SPA24.) This conclusion totally disregards existing and emerging markets for the uses at issue in this case. The District Court failed to consider the consequences its ruling will have on uses of copyrighted works by actors *other* than Google, and it ignores the reality of today's digital world, in which the question is not if, but when, a data breach will occur.

1. Google Destroyed Existing and Emerging Markets for Digital Books

It is well established that the "impact on potential licensing opportunities for traditional, reasonable or likely to be developed markets should be legally cognizable when evaluating a secondary use's 'effect upon the potential market for or value of the copyrighted work.'" *Texaco*, 60 F.3d at 930. The undisputed evidence establishes that Google's activities will harm the Authors by undermining existing and emerging licensing opportunities.

In *Texaco*, this Court considered whether a corporate library's systematic photocopying of scientific journal articles for research and archival purposes was fair use. 60 F.3d at 91. In analyzing whether it was appropriate for the district court to have considered the increase in publishers' revenue that would result if Texaco's unauthorized copying was not permitted as fair use, this Court observed that "the publishers still have not established a conventional market for the direct sale and distribution of individual articles[.]" *Id*. Nevertheless, the Court found that "they have created, primarily through [the Copyright Clearance Center], a workable market for institutional users to obtain licenses for the right to produce their own copies of individual articles via photocopying." *Id*. The Court reasoned that:

> it is not unsound to conclude that the right to seek payment for a particular use tends to become legally cognizable under the fourth fair use factor when the means for paying for such a use is made easier. This notion is not troubling: it is sensible that a particular unauthorized use should be considered "more fair" when there is no ready market or means to pay for the use, while such an unauthorized use should be considered "less fair" when there is a ready market or means to pay for the use.

Id. at 930-931.

Here, there is a "ready market or means" for Google's uses. A prime example is Google's own Partner Program, in which more than 45,000 rightsholders have given Google permission to display partial book content on the Internet in exchange for compensation from ad revenues. (CA92.) Beginning in

2004, the Partner Program allowed publishers and other rightsholders to permit Google to display their works in exchange for a split of ad revenue. Google could have extended this licensing model to meet the broader aims of the Google Books project. Indeed, Google *has* paid for the ability to use millions of copyrighted works; it just hasn't paid the Authors. Instead, it has opted to pay independent contractors and scanning vendors hundreds of millions of dollars, and it has paid libraries with the currency of infringing e-books.

Further, as described in detail in the expert report of Professor Daniel Gervais, collective management organizations provide a market-based mechanism by which libraries could compensate authors and rightsholders in exchange for a license to mass digitize and make various uses of copyrighted books in their collections. (*See* A785-87.) Just as in *Texaco*, the CCC and other collective management organizations around the world presently license the same types of materials that are scanned by Google, generating a significant stream of revenue for rightsholders. (A785, 791.) For example, authors and their associational representatives in Norway and Sweden have entered into or are finalizing license agreements to permit their national libraries to digitize entire collections of books in exchange for royalty payments.[11] In 2012, Google announced that in France "it

[11] *See* Alexis C. Madrigal, *Norway Decided to Digitize All the Norwegian Books*, The Atlantic, December 3, 2013, http://www.theatlantic.com/technology/archive/2013/12/norway-decided-to-

had an industrywide book-scanning agreement in place to cover works that are out of print but still under copyright—a category that covers most of the world's books," and that Google was "interested in exporting these deals elsewhere." Eric Pfanner, *Google Has Deal in France For Book-Scanning Project*, N.Y. Times, June 12, 2012, at B5.

Although the District Court ultimately rejected the Amended Settlement Agreement, the proposed resolution provides further compelling evidence that there is a ready market or means to pay for the uses that Google is making. Among other things, the ASA provided a mechanism to compensate the millions of authors whose copyrighted works had been digitized by Google without authorization. (A85-87, 116-26.) Under the ASA, the class of affected authors and rightsholders, including the Authors, would have granted a license to Google to digitize works and sometimes sell, display and make certain non-display uses of the works it had scanned, in exchange for a portion of the revenues earned from those uses. (A85-87, 112-13.) As part of this license, Google was expressly authorized to index the contents of the digitized works for search purposes and to allow researchers to conduct "non-consumptive research" using the digitized corpus (ASA, ECF No. 770-1 at § 7.2(d).) The ASA shows how a collective

---

digitize-all-the-norwegian-books/282008/; Jerker Rydén, *Sweden's Digital Library*, March 3, 2011, http://slidesha.re/1luSlna.

management system might work to permit Google's activities in this case while providing compensation to copyright owners.

In addition to thwarting the development of emerging and potential markets, Google's various unauthorized uses also undermine *existing* licenses for the use of the Authors' books. Rightsholders routinely grant online distributors a license to index their books and make them searchable as part of a commercial arrangement intended to promote book sales. (*See, e.g.*, A56; A67.) Amazon's Search Inside the Book program, described above, is a prime example. *See supra* at 2-3. Google's Library Project was designed to drive traffic away from Amazon, an online retailer, and toward Google. (CA 440). As courts have held, a secondary use that replaces a comparable service licensed by the copyright holder, even without charge, may cause market harm. *See Infinity*, 150 F.3d at 111 (although copyright holder provided a comparable service to its customers at no additional cost, defendant's use "replaces" the broadcaster "as the supplier of those broadcasts to meet the demand of his customers"); *ASCAP*, 599 F. Supp. 2d at 422, 432 (performance rights organization established existence of market for ringtone previews by showing that musicians generally granted a license to online distributors to allow their users to search for and play ringtone clips for promotional purposes).

2. The District Court Ignored Evidence of the Risk and Immense Consequences of a Data Breach

Google is playing with fire. Google's Library Project converts print books into digital files and then replicates those files, stores them on multiple servers connected to the Internet, and grants its Library Partners access to an online portal that permits them to simultaneously download millions of e-books onto their own computers. These unauthorized activities subject the Authors' books to the same types of risks of digital piracy that battered the market for sound recordings. While this issue was fully briefed below, the District Court failed to address it at all. (*Compare* Mem. Supp. Pls.' Mot. Summ. J. at 45-46, ECF No. 1050; Pls.' Reply Supp. Mot. Summ. J. at 16-17, ECF. No. 1085, *with* SPA24-28.)

The risks of data breaches are real and on the rise,[12] leading many information security professionals acknowledge that breaches are sure to occur. A day hardly passes without reports of damaging cyberattacks on even the most sophisticated technology companies, including Google. There are people who specifically target digital libraries and similar databases, and they have succeeded. In 2011, an activist was indicted for hacking into a proprietary database of journal articles by sneaking into a network interface closet in the MIT library, hooking his

[12] *See, e.g.*, 2013 Data Breach Investigations Report (2013), http://www.verizonenterprise.com/DBIR/2013/.

laptop directly into the network and downloading *over 4.8 million articles*, with the intent to disseminate the archive throughout the Internet. *See* Indictment, *U.S. v. Swartz*, No. 11-CR-10260 (D. Mass. July 14, 2011).

As it must, Google acknowledges there is a risk that its "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." (A562.) To make matters worse, Google does not monitor or control the security of the e-books it provides to its Library Partners (*see* CA508 ¶ 108), subjecting rightsholders to the potentially inadequate security measures of libraries and exacerbating the risk of a security breach (A802-03). In this way, Google's unauthorized uses create new risks to the literary market, further demonstrating that the Library Project is likely to cause harm to the "potential market" for the Authors' books.

Google, for its part, dismisses plaintiffs' security concerns as "doomsday hacker scenarios" (Google Opp'n Br. at 35, ECF No. 1072.) But there is nothing remote or speculative about such online security risks. *See, e.g.*, Richard Pérez-Peña, *Campuses Face Rising Threat from Hackers*, N.Y. Times, July 17, 2013, at A1; Wallace D. Loh, UMD Data Breach: Update #6 (Feb. 25, 2014), http://uhr.umd.edu/2014/02/umd-data-breach-update-6-02252014/ ("In the past couple of years, some 20 large universities across the country have [] reported major data

breaches. There is an arms race between hackers playing offense and universities playing defense.").

Nor is there anything remote or speculative about the fact that piracy can have devastating consequences for rightsholders. *See, e.g.*, S. Rep. No. 105-90, at 61 (1998) (recognizing "risk that uncontrolled public access to [libraries'] copies . . . in digital formats could substantially harm the interests of the copyright owner by facilitating immediate, flawless and widespread reproduction and distribution of additional copies . . . of the work"). Indeed, a 2010 study showed that publishers lost almost $3 billion in sales as a result of e-book piracy. Gini Graham Scott, *The Battle Against Internet Book Piracy* 12 (2013). Even if Google has thus far avoided a data breach, there is no assurance it will succeed in the future. Security could be compromised by increased pressure from hackers, foreign governments or private actors determined to disrupt American interests, budget cuts, a disgruntled employee or mistakes that lead to the release of copyright-protected materials.

The ASA reflected the importance of allowing rightsholders to decide whether and how to make their works available on the Internet. It included specific security procedures agreed upon by authors, a procedure for security audits and remedies in the event of any potential breaches. (*See* ASA, ECF No. 770-1, Art. VIII.) In sharp contrast, Google's unilateral mass digitization offers none of these protections.

Finally, it is the Authors, not Google, who should be entitled to decide whether "to reproduce" and "to authorize" others "to reproduce" their works into digital format, 17 U.S.C. § 106, and thereby subject their works to risks of piracy. *See Salinger v. Random House, Inc.*, 811 F.2d 90, 99 (2d Cir. 1987), (author "has the right to change his mind" regarding whether or not to publish work and is "entitled to protect his *opportunity* to sell his [works]"). By undermining Authors' rights to control whether, how and by whom their works are to be digitally exploited, Google disempowers Authors and diminishes the value of their works.

3. The District Court's Decision Clears the Way for Other, Less Responsible Parties to Engage in Their Own Mass Digitization Programs

In analyzing the fourth factor, a court must consider the consequences to the market if the infringer's conduct were to become "unrestricted and widespread." *Campbell*, 510 U.S. at 590; *see Texaco*, 60 F.3d at 927 n.12. If this Court upholds the District Court's determination that the Library Project constitutes fair use, it will open the door to others to engage in their own unauthorized digitization programs, further exacerbating the harm to the Authors and putting more books within the reach of digital thieves on countless more servers. Absent a legal obligation, these actors are unlikely to implement the security measures necessary to safeguard intellectual property that does not belong to them, especially when armed with a decision that the public interest in disseminating information easily

trumps the interests of copyright holders.[13] Even if these actors may want to implement necessary security measures, they may not have the technological or financial means to do so. As a result, the security risk posed by Google's uses, already multiplied by Google's distribution to its Library Partners, would be multiplied yet again.

For all these reasons, the fourth factor, like all of the factors in this fair use analysis, weighs against Google.

II.

A FUNDAMENTAL CHANGE IN THE BALANCE BETWEEN THE RIGHTS OF COPYRIGHT HOLDERS AND USERS MUST BE MADE BY CONGRESS, NOT BY A PRIVATE COMMERCIAL ENTITY

In determining whether Google's mass digitization program constitutes fair use, the Court should also take account of the fact that in drafting the Copyright Act, Congress recognized and granted rightsholders protections against the dangers inherent in copying protected works into digital format. As early as 1963, at a Copyright Office hearing, witnesses expressed concerns over the rise of computer technology and the rightsholders' loss of control that would result if users were permitted to copy books and other works into machine-readable format without

[13] *Accord Harper*, 471 U.S. at 559 ("[T]o propose that fair use be imposed whenever the social value [of dissemination] . . . outweighs any detriment to the artist, would be to propose depriving copyright owners of their right in the property precisely when they encounter those users who could afford to pay for it.") (alteration in original, internal quotation marks omitted).

authorization. These and other concerns led to the inclusion in Section 106 of the Copyright Act of language clarifying that rightsholders have the exclusive right "to do and *to authorize"* others to copy, display and distribute copyrighted works, including in digital format. *See* 3 *Omnibus Copyright Revision Legislative History* 120-27 (George Grossman ed., 2001).

Moreover, Section 108 of the Copyright Act, which gives libraries and archives the right to make copies of printed books in their collections under very limited circumstances (*see supra* note 6), was enacted after years of debate over the appropriate balance between rightsholders and users. It was not until the 1998 passage of the Digital Millennium Copyright Act that Section 108 was amended to permit libraries to make and use copies of books in digital format. However, the amended statute included important restrictions on libraries' ability to use digital copies, including that no more than three digital copies could be made and that digital copies not be "made available to the public in that format outside the premises of the library or archives in lawful possession of such copy." 17 U.S.C § 108(c)(2). Congress placed these restrictions on digital copying "in recognition of the risk that uncontrolled public access to the copies [] in digital formats could substantially harm the interests of the copyright owner by facilitating immediate, flawless and widespread reproduction and distribution of additional copies [] of the work." S. Rep. No. 105-90, at 61-62 (1998).

These very specific rights and limitations addressing libraries and archives make it doubly implausible to read the Act as authorizing a commercial entity like Google to expose copyrighted works to the very risk Congress sought to address. As the District Court noted in rejecting the ASA, if there is going to be a major change in the established balance between rightsholders and users, it should be made by Congress not Google. *See Authors Guild*, 770 F. Supp. 2d at 677.

CONCLUSION

Plaintiffs-Appellants respectfully request that this Court vacate the Judgment of the District Court; mandate that the District Court grant the Authors' motion for summary judgment and remand the case for consideration of a proper remedy that includes payment to Authors for the use of their works, an end to Google's provision of e-books to libraries and real protections from the risk of theft and dissemination.

Dated: New York, New York
April 7, 2014

FRANKFURT KURNIT KLEIN & SELZ, P.C.

By: /s/ Edward H. Rosenthal
Edward H. Rosenthal
Jeremy S. Goldman
Anna Kadyshevich
Andrew D. Jacobs
FRANKFURT KURNIT KLEIN & SELZ, P.C.
488 Madison Avenue, 10th Floor
New York, New York 10022
(212) 980-0120

Paul M. Smith
JENNER & BLOCK LLP
1099 New York Avenue NW, Suite 900
Washington, D.C. 20001
(202) 639-6000

Attorneys for Plaintiffs-Appellants

CERTIFICATE OF COMPLIANCE

This brief complies with the type-volume limitation of Fed. R. App. P. 32(a)(7)(B) because based on the word count of the word-processing system used to prepare the brief (Microsoft Word), this brief contains 13,583 words, excluding the part of the brief exempted by Fed. R. App. P. 32(a)(7)(B)(iii).

This brief complies with the typeface requirements of Fed. R. App. P. 32(a)(5) and the type style requirements of Fed. R. App. P. 32(a)(6) because this brief has been prepared in a proportionally spaced typeface using Microsoft Word in 14-point Times New Roman.

Dated: April 7, 2014

By: /s/ Edward H. Rosenthal
Edward H. Rosenthal