October 27, 2015

Weil

Alert Intellectual Property/Media

Second Circuit Finds That Google's Scanning of Full Text and Display of "Snippets" Are Fair Uses of Copyrighted Books

By Jonathan Bloom and Reed Collins

Introduction

In a highly anticipated and not unexpected ruling, the Second Circuit affirmed the district court's decision on summary judgment that Google engaged in fair use when it digitally copied millions of books, including books still protected by copyright, without permission in order to create a searchable database that returned short, one-eighth of a page "snippets" of text in response to keyword searches. Drawing a distinction between exploitation of a work's expressive content and Google's dissemination of "information *about*" the work, the court held that Google's copying served a transformative purpose and did not harm what the court found to be the different market for licenses to access substantial portions of the books. Following the court's June 2014 ruling in the factually related case *Authors Guild, Inc. v. HathiTrust*,[1] the court's second evaluation of Google book-scanning confirmed the centrality of transformativeness to fair-use analysis as well as the increased breadth of that concept in the Internet Age, while underscoring the more limited role of the commercial purpose of Google's use. *Authors Guild v. Google, Inc.*, No. 13-4829-cv, 2015 WL 6079426 (2d Cir. Oct. 16, 2015).

Google carries the substantial added weight of having been written by Judge Pierre N. Leval, whose influential 1990 law review article on fair use[2] articulated the concept of transformative use that was adopted by the Supreme Court in *Campbell v. Acuff-Rose Music, Inc.*,[3] the Court's most recent fair-use ruling. In assessing how *Campbell* and its progeny applied to the Google Books and Google Library projects, Judge Leval was effectively passing judgment on whether the concept of transformative value had evolved in keeping with his seminal formulation of it. With Judge Leval's embrace in *Google* of the line of post-*Campbell* cases that have rejected claims involving search engines and other "functional" new uses of copyrighted works, fair use becomes further entrenched as a viable defense for the copying required to create innovative digital information tools and for limited, non-substitutional access to copyrighted works provided by such tools.

The *Google* decision is also notable for its criticism (both implicit and explicit) of rulings by other courts of appeal that, in the Second Circuit's view, misconstrued aspects of fair use. These include the Eleventh Circuit's ruling giving greater weight to the nonprofit educational purpose of digital scanning for course readings than to its lack of transformative value[4] and the Seventh Circuit's failure in *Kienitz v. Sconnie Nation LLC* to distinguish clearly between derivative works and transformative works.[5]

Factual Background

Google has made digital copies of more than twenty million books (some copyrighted, some public domain, most non-fiction and mostly out of print) submitted to it by major libraries through its Google Library Project, without the permission of rights holders. The Google Books search function allows Internet users to search the digital copies for a list of all books in the database in which the search terms appear, the number of times the term appears, a brief description of each book, and, sometimes, links to buy the book online or to libraries where the book can be found.[6] Google's "ngrams" research tool also furnishes statistical information about the frequency of word and phrase usage over time and space, allowing users to glean information about "syntactic patterns and thematic markers."[7]

A Google Books search further allows users a limited viewing of the text of each book by displaying a maximum of three "snippets" of text (an eighth of a formatted page, *i.e.* 3 lines of a 24-line page) containing the searched-for keywords. The same three snippets are returned for every search for a particular word or term no matter how many times it is searched, and only the first usage of a term on a given page is displayed. One snippet per page and one complete page out of every ten pages is "blacklisted" and completely unavailable for viewing under any circumstances (about 22% of total content). No snippets are available for dictionaries, cookbooks, books of short poems, and other books for which a single snippet is likely to satisfy the need for the book. Rights holders may exclude their books from snippet view completely by submitting a request to Google.[8]

No advertising is displayed to users, and Google does not receive any payment if a searcher uses a link to buy a book.[9] Google allows the participating libraries to download and retain the digital copies of books they submit under agreements that commit the libraries not to use their digital copies in violation of copyright law.[10]

The plaintiffs whose claims reached the Second Circuit were three book authors who sued Google for copyright infringement, seeking injunctive relief, declaratory relief, and damages.[11]

The District Court Ruling

Judge Denny Chin, who retained the case after being elevated to the Second Circuit, granted summary judgment for Google, holding that the Google Books and Google Library projects constituted fair use.[12] Judge Chin found Google's use of the works to be "highly transformative," turning "expressive text into a comprehensive word index that helps readers, scholars, researchers, and others find books."[13] He further held that Google's display of snippets was properly limited and that, far from substituting for the original works in the marketplace, "there can be no doubt but that Google Books improves book sales."[14]

The Second Circuit Ruling

The Second Circuit, in a 46-page opinion written by Judge Leval, affirmed. The decision offers both general insights on the fair use doctrine and specific guidance as to each of the fair use factors set forth in section 107 of the Copyright Act. The court described the statutory factors as guideposts toward answering 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."[15] The court's approach to the four factors positioned transformative value at the center of the analysis.

As to the first factor, the "purpose and character of the use," the court began with the observation in *Campbell* that "the creation of transformative works . . . lie[s] at the heart of the fair use doctrine's guarantee of breathing space within the confines of copyright."[16] Distinguishing the concept of a derivative work, which primarily involves "changes of form" (such as the adaptation of a novel into a screenplay), the court explained that 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."[17] By that standard, building a searchable database from millions of printed books was a "quintessentially transformative" purpose, as the court had previously recognized in *HathiTrust*.[18] Although the searches in *HathiTrust* did not display "snippets," the court held

that Google's snippets likewise contributed to the "highly transformative purpose of identifying books of interest to the searcher."[19] The court cited approvingly a line of circuit court rulings in which uses involving the "provision of information about" copyrighted works, including by search engines, were held to be fair use.[20]

The statute further directs courts to consider whether the use is commercial or nonprofit in nature, but the court placed greater weight on transformative use, noting that the Second Circuit has "repeatedly rejected the contention that commercial motivation should outweigh a convincing transformative purpose and absence of significant substitutive competition with the original."[21] (The court pointed out that there also is no presumption that a nonprofit educational purpose qualifies as fair use. Surely alluding to the Eleventh Circuit's recent "digital coursepack" decision in *Cambridge University Press v. Patton*,[22] the court stated that "[a]uthors who write for educational purposes, and publishers who invest substantial funds to 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."[23]) Finding no direct competition between Google's database offering and the original books (as discussed in factor four below), the court was unpersuaded by plaintiffs' contention that Google's ultimate motive to profit from "dominance of the world-wide Internet search market" should preclude a finding of fair use.[24]

The second factor – the nature of the copyrighted work – is, the court noted, rarely dispositive.[25] Although the Supreme Court has stated that the law "generally recognizes a greater need to disseminate factual works than works of fiction or fantasy," the *Google* court referred to that statement as a "passing observation in dictum[26] and observed that authors of factual works are entitled to copyright protection for their "manner of expressing those facts and ideas."[27] Thus, while the plaintiffs' books in this case were factual, the court was not influenced "one way or the other" by factor two.[28]

As to the third factor – the "amount and substantiality of the portion used in relation to the copyrighted work as a whole" – the court emphasized *Campbell*'s directive that the extent of permissible copying "varies with the purpose and character of the use."[29] Here, Google's transformative purpose justified its extensive copying, given that, as in *HathiTrust*, copying the entirety of each book was "literally necessary" to create a full-text search function.[30] As for Google's snippet view, the court held that the relevant inquiry was not the amount of the work copied (the entire book), but rather the amount of the work "made accessible" and thus potentially a substitute for the original (very little).[31] Specifically, Google had taken steps to ensure that, among other things: (a) snippets were only an eighth of a page and arbitrarily and uniformly divided by lines of text (not complete sentences, paragraphs, or any measure dictated by content); (b) only three snippets were shown per search in a random order; (c) the same three snippets were always shown for any given keyword, preventing users from "gaming the system" through repeated searches; (d) some snippets and whole pages were "blacklisted" to prevent completeness in search results; and (e) no snippets were available from dictionaries or other types of books in which a small segment was likely to satisfy the searcher's need for the content.[32] Although the plaintiffs hired researchers to run multiple searches over a period of weeks, they were unable to access even sixteen percent of text, and the portions they *could* access were not sequential but "scattered randomly throughout book."[33]

The court emphasized that factor four — the "effect of the [copying] use upon the potential market for the value of the copyrighted work" — is linked closely to factor one, as transformative uses by their nature are less likely to serve as "a satisfactory substitute for the original."[34] Still, the court explained that the fourth factor demanded analysis of whether, notwithstanding its transformative purpose, snippet view harmed the value of the copyrighted originals by providing a market substitute.[35] As to this issue, the court held that Google's snippet function was not an "effectively competing substitute for Plaintiffs' books" in light of the limiting measures listed above, which made any attempt to aggregate the snippets "cumbersome, disjointed, and incomplete."[36] While recognizing that snippets could "cause *some* loss of sales," the court

could see no "meaningful or significant" effect upon the market for the original works.[37] Careful to limit the scope of the ruling, however, the court cautioned that an arrangement that would allow public users to read "substantial portions of the book … would most likely constitute copyright infringement if not licensed by the rights holders."[38]

The court went on to address – and reject – several of the plaintiffs' specific arguments.

First, the plaintiffs argued that even if Google's activities did not impair any market for the original books, Google nonetheless infringed their *derivative rights* in search functions, depriving them of revenue they could obtain by licensing their works to search providers. But the court held that authors' copyrights "do[] not include an exclusive right to furnish the kind of *information about* the works that Google's programs provide to the public."[39] Rather, the exclusive right to prepare derivative works extends only to "re-present[ing] the protected aspects of a work, *i.e.* its expressive content," in an "altered form."[40] The court distinguished Google's snippets, which it found do not provide "any meaningful experience of the expressive content of the book," from ringtones, which are licensable derivative works because they "provide[] a mini-performance of the most appealing segment of the authors expressive content."[41]

Second, the plaintiffs argued that Google's storage of digital copies of their books was not fair use because it created an unreasonable risk that hackers would access and distribute the copies, destroying the value of their copyrights. Calling this argument "theoretically sound," the court rejected it as a factual matter because Google's "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."[42] The plaintiffs failed to rebut Google's "sufficient showing of protection of its digitized copies of Plaintiffs' works."[43]

Finally, the plaintiffs contended that by distributing complete digital copies to the libraries that had supplied the printed books, Google deprived them of the opportunity to license the libraries' digital access to their works. But the court held that this too was fair use and no different than if the libraries had created their own digital copies for search purposes – a practice the court previously approved in *HathiTrust*.[44] Although the plaintiffs emphasized the risk that the libraries might make infringing uses of Google's digital copies, the court noted that Google contractually required each participating library to use the copies only in non-infringing ways and to take precautions to prevent dissemination of the copies to the public at large.[45] The court observed that were there any "evidence that Google was aware of or encouraged" infringing practices by the libraries, Google "could be liable as a contributory infringer."[46]

Conclusion

Building on last year's decision in *HathiTrust*, the Second Circuit has now clarified that even copying of an entire work by a for-profit corporation can be fair use so long as the copying, and the uses the defendant makes of the copies, advance public knowledge in such a way as not to cause market substitution for the original works. The record evidence of the specific steps Google took to limit the amount of copied expression made accessible to the public through the Google Library Project – steps, the court found, that minimized any risk of market harm to the authors and reflected the different purpose for which Google was using the books it copied – were enough to overcome the plaintiffs' concerns. The court held that the transformative (as opposed to derivative) nature of Google's use, with its value to linguistic researchers and others, placed it outside the scope of the plaintiffs' licensing rights. Importantly, however, the court warned that allowing public access to "substantial portions of a book" without permission would likely be infringing.

1. 755 F.3d 87 (2d Cir. 2014).
2. Pierre N. Leval, *Toward a Fair Use Standard*, 103 Harv. L. Rev. 1105 (1990).
3. 510 U.S. 569 (1994).
4. *See Cambridge University Press v. Patton*, 769 F.3d 1232 (11th Cir. 2014).
5. 766 F.3d 756 (7th Cir. 2014), *cert denied*, 135 S. Ct. 1555 (2015).
6. *Google*, 2015 WL 6079426, at *2-3.
7. *Id.* at *3.

8. *Id.* at *3-4, *14.
9. *Id.* at *3.
10. *Id.* at *4.
11. The Authors Guild, a membership organization of published authors, was also a plaintiff-appellant, but the Second Circuit previously held that the Authors Guild lacked standing to sue for copyright infringement on its members' behalf. *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 94 (2d Cir. 2014), *cited in Google*, 2015 WL 6079426, at *2 n.1.
12. *Authors Guild, Inc. v. Google, Inc.*, 954 F. Supp. 2d 282, 294 (S.D.N.Y. 2013).
13. *Id.* at 291.
14. *Id.* at 292-93.
15. *Google*, 2015 WL 6079426, at *6.
16. *Id.* at *7 (quoting *Campbell*, 510 U.S. at 579).
17. *Id.* at *7-8.
18. *Id.* at *9 (quoting *HathiTrust,* 755 F.3d at 97).
19. *Id.* at *10.
20. *Id.* at *8 & n.17.
21. *Id.* at *11 (citing *Carious v. Prince*, 714 F.3d 694, 708 (2d Cir. 2013), and *Castle Rock Entm't, Inc. v. Carol Pub. Grp., Inc.*, 150 F.3d 132, 141-42 (2d Cir. 1998)).
22. 769 F.3d 1232 (11th Cir. 2014).
23. *Google*, 2015 WL 6079426, at *11 n. 20.
24. *Id.* at *1.
25. *Id.* at *11.
26. *Id.* (quoting *Harper & Row Pubs., Inc. v. Nation Enters.*, 471 U.S. 539, 563 (1985)).
27. *Id.* at *11.
28. *Id.* at *12.
29. *Google*, 2015 WL 6079426, at *12 (quoting *Campbell*, 510 U.S. at 586-87).
30. *Id.* at *13.
31. *Id.*
32. *Id.* at 13-14.
33. *Id.* at *14.
34. *Id.* at *15.
35. *Google*, 2015 WL 6079426, at *15
36. *Id.* at *16.
37. *Id.* at *15.
38. *Id.* at *17
39. *Id.* at *16 (emphasis added).
40. *Id.*
41. *Google*, 2015 WL 6079426, at *17 (discussing *Infinity Broad. Corp. v. Kirkwood*, 150 F.3d 104 (2d Cir. 1998), and *United States v. ASCAP*, 599 F. Supp. 2d 415 (S.D.N.Y. 2009)).
42. *Id.* at *18.
43. *Id.*
44. *Id.* at *19.
45. *Id.*
46. *Id.*

If you have questions concerning the contents of this issue, or would like more information about Weil's IP/Media practice group, please speak to your regular contact at Weil, or to the editors or practice group members listed below:

Editors:

Randi Singer (NY)	Bio Page	randi.singer@weil.com	+1 212 310 8152
Jonathan Bloom (NY)	Bio Page	jonathan.bloom@weil.com	+1 212 310 8775

Contributing Authors:

Jonathan Bloom (NY)	Bio Page	jonathan.bloom@weil.com	+1 212 310 8775
Reed Collins (NY)	Bio Page	reed.collins@weil.com	+1 212 310 8497