

Google Books and Digitisation of Libraries: Fair Use or Extension of Copyright?

Mary G. Mayiladumpara

Introduction

Initially known as Google Print, Google Book Search is a tool from Google Inc., the Silicon Valley search giant, which searches the full text of books that are scanned and stored in Google's digital database. The books are scanned by using an undisclosed proprietary method. Google Book Search consists of a partner program with publishers called Google Book Search Publisher Program and an effort by Google in conjunction with several libraries to scan books into a digital format called Google Book Search Library Project.

Under the Publisher Program the copyright owners authorise Google to scan the full text of the book into Google's search database. Since there is an agreement between Google and the copyright holder, and advertising revenue earned by Google is shared with publishers, the Publisher Program does not raise copyright issues.

Under the Library Project, Google scans into its search database books from the libraries of Harvard, Stanford, and Oxford Universities, the University of Michigan, and the New York Public Library. In response to search queries by users of Google search engine or Google Book Search, users will be able to browse the full text of public domain materials, but only a few sentences of text around the search term in books still covered by copyright. The copying of public domain books does not lead to any copyright issues. The whole controversy surrounding Google Book Search relates to scanning full text of books still under copyright.

Google Book Search Library Project– the *modus operandi* and legal issues

The Library Project involves two steps that raise copyright issues: full text scanning of in-copyright books and display of snippets from the scanned book in response to user queries. The most vital aspect in Google Book Search of digitised in-copyright books scanned as part of the Library Project is that not more than three "snippets" are displayed during a search. A "snippet" displays a few sentences on either side of the search term. Neither the full text nor even a full page is displayed to a user in case of in-copyright books scanned as part of the Library Project. Google asserts that the digitisation of books under copyright as part of the Library Project and its use in Google Book Search constitutes a "fair use" of the work.

Publishers are provided with two options in case of an in-copyright book digitised under the Library Project: they may either transfer a copyrighted book scanned under the Library Project to the Publisher Program or opt their in-copyright books out of the project. The opt out mechanism has been criticised as being contrary to the traditional notions of copyright. Patricia Schroeder, President of Association of American Publishers has asserted that Google's opt out procedure "shifts the responsibility for preventing

infringement to the copyright owner rather than the user, turning every principle of copyright law on its ear.”¹ Google essentially contends that its opt out program negates any infringement liability. But, if infringement were found, Google argues that its activity is protected by copyright’s fair use doctrine.²

“Opt out” Program – an Extension of Copyright?

The general rule of copyright law requires a prospective user to seek permission for use; Google has reversed the rule through the opt out mechanism whereby entire collections of the contributing libraries will be scanned and indexed by Google *unless* a copyright owner opts out.³ However, Google has stated that actively seeking out rights’ holders to obtain their permission involves prohibitive transactional costs and is impracticable. It is Google’s contention that the prohibitive transactional cost of obtaining copyright owners’ consent entitles Google to a fair use right that permits it to digitise copyrighted works unless the rights’ owner denies permission.

The Publishers claim that Google’s opt out program “stands copyright law on its head.”⁴ They argue that one cannot generally announce one’s intention to infringe multiple copyrighted works and collectively offer copyright holders the opportunity not to have their work infringed. The opt out method may seem contrary to copyright law since the one who copies or distributes a copyrighted work bears the burden of requesting permission from a copyright holder. However, when search engines crawl the web and index content, they do not formally request permission for such use despite the fact that such indexes copy entire webpages. Internet has built-in mechanisms to protect contents on the web from being indexed by search engines. The opt out program may thus be considered a necessary extension of traditional copyright principles to meet the challenges of a digital world.

The Fair Use Defence

It is a settled principle of copyright law that not every use of a copyrighted work constitutes infringement. The exclusive rights accorded to a copyright owner under 17 USC § 106 are “to do and to authorize” any of the activities specified in the five numbered clauses therein are subject to the limitation of “fair use” under 17 USC § 107. The affirmative defense of fair use, permits copying in certain circumstances, particularly where value is being added to a copyrighted work.

There are two approaches to the application of the fair use defence to digitisation of in-copyright books by Google as part of the Library Program:

¹ Association of American Publishers Press Release, “Google Library Project Raises Serious Questions for Publishers and Authors,” August 12, 2005.

² Congressional Report Service: The Google Book Search Project: Is Online Indexing a Fair Use Under Copyright Law? available at http://openers.com/rpts/RS22356_20051228.pdf.

³ *Ibid* at p. CRS 2

⁴ *Publishers: Value of Book Search Project Shows That Scanning Is Not Fair Use*, 71 BNA PATENT, TRADEMARK & COPYRIGHT J. 94 (Nov. 25. 2005)

- (i) that intermediary copying of full texts of in-copyright books for the purpose of providing snippets in response to users' search queries is a fair use in itself:

There is no obligation on a user to obtain the consent of the copyright owner for a "fair use" of the work. This would mean that that Google will not be required to remove scanned in-copyright books from its database even if rights' holders deny their consent, provided Google does not provide the full text to a user.

- (ii) that intermediary copying of full texts of in-copyright books for the purpose of providing snippets in response to users' search queries is a fair use as long as copyright holders do not opt out of the program:

Due to the prohibitive transactional cost of obtaining consent from rights' holders, copying full texts of in-copyright work will constitute "fair use" as long as copyright holders do not deny their consent or "opt out".

Google had initially proceeded on the basis of the first approach. When publishers questioned the first approach, Google announced (on August 11, 2005) suspension of the Library Project (until November) so that publishers could decide whether they want to opt their in-copyright books out of the project; thereby adopting the second approach.

Analysis of Google's Fair Use Defence – Weighing the Four Factors.

17 USC § 107 lays down the four-pronged "fair use" test which was recognised as early as in nineteenth century opinion by Justice Story in *Folsom v. Marsh*⁵ wherein he laid out an approach for analyzing a question of fair use by looking into "the nature and objects of the selections made, the quantity and value of the materials used, and the degree in which the use may prejudice the sale, or diminish the profits, or supersede the objects of the original work."⁶ The four factors test laid down in 17 USC § 107 specifies (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.

(i) The purpose and character of the Use

The first "fair use" factor set forth in section 107 of Title 17 of the United States Code is "[t]he purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes." As the statutory language indicates, of particular importance here is whether the use at issue is commercial. However, the commercial nature of a use is not determinative.⁷ In considering the "purpose and character" of the use under 17 USC § 107, courts have relied on whether a work is "transformative,"⁸ that is, whether the use "adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message." In *Campbell v. Acuff-Rose Music, Inc.*, the Supreme Court examined the first statutory

⁵ 9 F. Cas. 342 (C.C.D. Mass. 1841)

⁶ *Ibid* at p. 348

⁷ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 584 (1994).

⁸ Pierre N. Leval, *Toward a Fair Use Standard*, 103 Harv. L. Rev. 1105, 1111 (1990)

factor and concluded that, "the more transformative the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use."

In *Kelly v. Arriba-Soft*,⁹ the Ninth Circuit considered legality of a search engine's practice of creating small reproductions ("thumbnails") of images and placing them on its own website (known as "inlining"). It was held to be a fair use since thumbnails were much smaller and of much poorer quality than the original images and served as an index of the images and helped the public access them although the full image was copied for the purpose of creating thumbnails.

Although Google's use is commercial, it is highly transformative. Google is effectively scanning the books and turning them into the world's most advanced searchable digital index. This makes Google's indexing of in-copyright similar to the case of Arriba Soft.

(ii) Nature of the Copyrighted Work

Under the second factor, courts examine whether a copyrighted work is factual or creative, and creative works are given greater protection.¹⁰ The books scanned and indexed by Google will be a mix of creative and factual works. Since Google scans millions of books ranging from creative to fact-based, the works need to be considered collectively.¹¹ A collective consideration of creative and factual works does not favor Google's fair use case. The fact that the books may not be published on the internet is likely to be irrelevant. The unpublished nature of a copied work may affect a court's fair use analysis.¹² All books scanned by Google being published books, this factor is also irrelevant.

(iii) The Amount and Substantiality of the Portion Used

The third statutory factor in the fair use test in 17 USC § 107 is "the amount and substantiality of the portion used in relation to the copyrighted work as a whole." In *Kelly v. Arriba*, although the Ninth Circuit recognized that "copying an entire work militates against a finding of fair use" the court nonetheless stated that "the extent of permissible copying varies with the purpose and character of the use" and proceeded to observe that "if the secondary user only copies as much as is necessary for his or her intended use, then this factor will not weigh against him or her." The Ninth Circuit further observed that "although Arriba did copy each of Kelly's images as a whole, *it was reasonable to do so in light of Arriba's use of the images*. It was necessary for Arriba to copy the entire image to allow users to recognize the image and decide whether to pursue more information about the image or the originating web site."

⁹ 03 C.D.O.S. 5888 (9th Cir. 2003)

¹⁰ *Campbell*

¹¹ Emily Anne Proskine, *Google's Technicolor Dreamcoat: A Copyright Analysis of the Google Book Search Library Project*, 21 Berkeley Tech. L.J. 213

¹² *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 564 (1985).

Scanning the entire text of books into Google's database is essential for the purpose of searching the entire text to provide relevant results to users and is therefore reasonable for the purpose of the effective operation of Google Book Search. Partial copying of text would defeat the very purpose of the use, as it would lead to incomplete results. Although the full text is scanned, the user views only "snippets". In order to cover the numerous possible search queries, numerous snippets have to be scanned, which, in effect implies the whole text. Unlike *Arriba*, which provided small reproductions of whole images (thumbnails), Google will not provide users with a copy of the entire work, but a few sentences on either side of the search term.

(iv) The effect of the Use on the Markets

In *Harper & Row*,¹³ the U.S. Supreme Court indicated that the fourth fair use factor is "undoubtedly the single most important element of fair use." This factor requires determination of the potential effect of the unauthorized use on the market for the original work. Although there is a presumption of market harm when a use is commercial, the presumption applies only to instances of "mere duplication for commercial purposes."¹⁴ A transformative use, even if commercial, does not warrant a presumption of market harm.¹⁵

In *Kelly v. Arriba-Soft*, the Ninth Circuit observed that reproductions of images in the form of "thumbnails" and placing them on a website did not undermine the potential market for the sale or licensing of those images. In *Sony Corporation of America v. Universal City Studios, Inc.*,¹⁶ the U.S. Supreme Court held that the videotaping of television broadcasts for private home use is a permissible fair use by relying on the district court's factual findings indicating that "the average member of the public uses a VTR principally to record a program he cannot view as it is being televised and then to watch it once at a later time (time-shifting.) Although the first three fair use factors were against a finding of fair use, the Court focused on the fourth factor. It emphasized that "a use that has no demonstrable effect upon the potential market for, or the value of, the copyrighted work need not be prohibited in order to protect the author's incentive to create."¹⁷

The publishers have contended that widespread digital copies of books will lead to less number of printed copies sold. The contrary is true. Since Google limits the search results to a few sentences before and after the search term, the Library Project will not conflict with the normal commercial exploitation of works. In *Kelly v. Arriba-Soft* the Ninth Circuit decided that "the effect of the use upon the potential market for or value of the copyrighted work," weighed in favour of *Arriba*. The court found that the *Arriba*'s "search engine would guide users to *Kelly*'s web site rather than away from it."¹⁸ The

¹³ 471 U.S. 539, at 564

¹⁴ *Campbell*, 510 U.S. at 591.

¹⁵ *Ibid.*; *Kelly*, 336 F.3d at 818-22

¹⁶ 464 U.S. 417 (1984).

¹⁷ *Ibid* at 450

¹⁸ *Kelly*, 336 F.3d at 821

court also observed that the thumbnail images would not harm Kelly's ability to sell or license full size images because the low resolution of the thumbnails effectively prevented their enlargement. The fourth test is more in favour of Google than Arriba since Google displays only snippets from the in-copyright book and the users are provided with links or details to online books stores and libraries where the particular book is available.

“Public Interest” – the fifth factor?

Since the underlying constitutional purpose of copyright law is to promote broad public availability of literature, music, and other arts,¹⁹ it needs to be considered whether the public interest promoted by Library Project should weigh in favour of Google's fair use defence. In *Sony Corp. of America v. Universal City Studios* the US Supreme Court observed that because fair use is an “equitable rule of reason” to be applied in light of the overall purposes of the Copyright Act, other relevant factors may also be considered.²⁰ The fair use doctrine serves the public interest in the reception and dissemination of information.²¹ Thus, public interest would surely qualify as a fair use factor.

In *Online Policy Group v. Diebold, Inc.*,²² the court it was held that the reproduction of an entire email communication without any kind of modification could still be considered fair use, where public interest is overwhelming. The emails in question, discussed potential defects in electronic voting machines, which clearly involved public interest.²³ In this case publication of the emails was characterised as a fair use in light of the public interest in legitimate elections and absence of any profit motive.

The Library Project definitely promotes public interest by making information available for search and by facilitating research. Mary Sue Coleman, in an address to the Professional/Scholarly Publishing Division of the Association of American Publishers (AAP) in Washington, D.C. aptly stated the benefits of digitising libraries by saying, "the University of Michigan's partnership with Google offers three overarching qualities that help fulfill our mission: the preservation of books; worldwide access to information; and, most importantly, the public good of the diffusion of knowledge”

The Fair Use defence in the United Kingdom

Given the fact that the Bodleian Library at Oxford University is one of libraries participating in Google Book Search, it is relevant to consider copyright law in the United Kingdom. United Kingdom's copyright law would be applicable if the physical scanning of books under copyright takes place in the U.K., or the resulting copies are

¹⁹ Bruce P. Keller & Jeffrey P. Cunard, *Copyright Law: A Practitioner's Guide* (2006) Practising Law Institute

²⁰ *Sony Corp.* at 448, 454

²¹ *Iowa State Univ. Research Found. v. Am. Broad. Cos., Inc.*, 621 F.2d 57, 60 (2d Cir. 1980).

²² 337 F. Supp. 2d 1195, 1203 (N.D. Cal. 2004).

²³ *Ibid* at 1203-1204

stored on servers located in the U.K.²⁴ Unlike the inclusive, flexible and potentially adaptive fair use exception under U.S. law, U.K. copyright law enumerates detailed and specific exceptions often designed with specific institutions and purposes in mind. Although “fairness” in a fair dealing assessment considers many of the factors similar to those in U.S. law, these factors are relevant only if the purpose of the use falls within articulated parameters of the specific institutions and purposes. None of the fair dealing categories in UK copyright law help Google’s case, as using a searchable electronic index of works for 'informational' purposes is simply not an exempted category of use.

Conclusion

The primary purpose of Copyright is to protect the commercial interest in works and also to protect the moral rights of the creator of any work. The Google Book Search Library Project does not prejudice the commercial interests of copyright owners; it also does not impinge the moral rights of authors. The only area where it militates against traditional notions of copyright is the scanning of entire text of books. This however is necessary to enable a comprehensive search. The use of copyrighted material in Library Project when looked at in the present day context, can safely be said to come within the ambit of fair use. However, to the extent that it requires owners of copyright to opt out of the project, it extends traditionally accepted principles. The ability to assert fair use will also depend on the jurisdiction within which the Google Book Search Library Project operates.

²⁴ Paul Ganley, *Google Book Search: Fair Use, Fair Dealing and the Case for Intermediary Copying*, (Working Paper, 13 January 2006)