8-2011

The Google Conundrum: Perpetrator or Facilitator on the Net? - Forging a Fair Copyright Framework of Rights, Liability and Responsibility in Response to Search Engine 2.0 - Part II: The Google Books Search Project

Warren B. CHIK

Singapore Management University, warrenchik@smu.edu.sg

DOI: https://doi.org/10.1016/j.clsr.2011.03.012

Published in Computer Law and Security Review, 2011 Aug, 27 (4), pp. 331–347.
http://doi.org/10.1016/j.clsr.2011.03.012

The Google conundrum: Perpetrator or facilitator on the net? - Forging a fair copyright framework of rights, liability and responsibility in response to search engine 2.0 — Part II: The Google Books Search Project

Warren B. Chik

Singapore Management University, Singapore

Keywords:
Google
Google Books Search Project
Google video search engine
Internet intermediaries regulation
Safe harbor
DMCA
Fair use

ABSTRACT

Is Google in its quest for search engine optimization through the creation of new technologies, which not only improves its search algorithms but also refines its search functions for users, doing it in a manner that makes it a perpetrator of primary copyright infringement or an invaluable facilitator for Internet functionality? How should the balance of interests in the treatment of creative works be recalibrated in the face of changes in search engine technology and operations, and the disputes that have arisen within the last decade in the context of the digital age and its needs? Using Google as a case study, this paper will look at the two main areas of dispute over the operations of information locator tools and services that either threatens search engine functionality and efficiency or weakens copyright holders' exclusive rights. It proposes a concerted set of solutions through a reassessment and amendment of copyright law to optimize the social benefits and objectives of both the copyright regime and technological innovations in the electronic model of information archiving, indexing and delivery. A fair distribution of responsibilities and allocation of rights and liabilities will be suggested. In the process, due consideration will be given to both public and private interests, with the former taking precedence; while the recommended solutions will be made within the currently outdated framework for Internet intermediary protection (i.e. safe harbor laws) and exceptions (i.e. specific statutory exemptions and the general fair use defense) under the existing copyright regime. Thus, the proposed changes will be far reaching without being too radical a departure from current law, an evolution that will likely be more acceptable and realistic a solution to the problem.

This paper is published in two parts. Part One of this paper published in the previous edition of the CLSR at [2011] 27 CLSR 110–131 dealt with the challenges to the copyright regime posed by the operations and technology behind the Google Images Search Engine, while Part Two will assess the benefits of the Google Books Search Project vis-à-vis the effects it will have on the scope of copyright protection. Recommendations are made to copyright law to accommodate both functions while generally preserving the main objectives of copyright protection.

1. Introduction

Due to the various objectives of a modern search engine, including its role as archiver, advertising agent, technology innovator and search market optimizer, and because of the idiosyncrasies of each type or form of work, in particular those that manifest in the written, image and audio-visual form; the manner in which technology is developed and employed to serve Google's role, function and operations varies and gives rise to complications that the existing law does no adequately or satisfactorily address. To put it in the context of the law, one main reason for the divide between law and technology lies in the fact that statutory immunity for Internet intermediaries that would include search engines is fundamentally based on a lack of control over third party material (i.e. enabling secondary infringement); whereas Google is increasingly involved in various aspects and stages of development and delivery of third party material irrespective of source and legality and have in certain cases and in the process and various stages of its services itself copied, edited and adapted works and materials (i.e. perpetrating primary infringement). Another reason is the conflict between societal and private interests involved in the copyright equation; that is, between default protection for copyright owners and legitimate user exemptions, which are given a new dimension due to the relatively new forms and uses that works are put to in the digital environment (e.g. User Generated Content (UGC)). The laws relating to Internet intermediaries have to some extent addressed the rights, liabilities and responsibilities of information locator tools. However, given the changes since, and even during, the time of the creation of these laws, they have left many issues and questions still open and unresolved to a large extent, particularly those that have arisen after the laws have already been passed statutorily.[1] Given the lag time between laws and technological changes, it is not surprising that it is left to the disputants and the courts to grapple with the issues in the interim period.

A compromise legal solution is necessary because the law and policy-makers can neither abandon the objectives of the copyright framework and risk harming the creative drive and multimedia industry nor the laws preserving the Internet and the WWW that would set the world back from socio-economic and technological progress facilitated by the digitization of both medium and media.

In Part One of this paper, I had set out the relationship between emerging technologies and the copyright regime, in particular, the increasingly complex and diversified web of search engine functions on the World Wide Web (WWW) with their differential treatment of creative content in pursuit of optimizing search engine functionality. Google Inc. was introduced as the subject matter of discussion due to the fact that it is the market leader in the business of search engines and performs most of the controversial functions that have given rise to disputes with various copyright holders. In fact, it is at the forefront of developing and operating search engine-related technologies. It has also been the main target for threats of legal action and the main protagonist in instituted legal proceedings on the issues relating to search functions and operations, which is the focus of this article.

While the focus of Part One was on the issues surrounding the functions of the Google Images Search Engine (GIS), the scope of Part Two will remain on the Google search engine but specifically on the controversies generated and disputes arisen with the creative industry in relation to the Google Books Search Project (GBS). The main policy considerations and some of the solutions recommended for the GIS that are also applicable to the GBS will be reiterated here. However, the GBS will also have its own set of problems and proposed solutions to meet some of its unique functions and objectives.

2. The objectives of this paper: the search for a compromise

The focus of this article is on the legal challenges that Google faces in two main areas of its operations relating to the expansion of its search engine functions, with a special focus on the technological tools that it had developed and is currently using to improve its information location services. In Part One of this paper, I have examined the US cases on the practices and features relating to the effective functioning of image search engines and made some recommendations for legal reform. In Section I of this Part, I will examine the features of the GBS that sets it apart from Google's other information search services (i.e. News, Finance, Scholar and Blogs) as well as briefly consider Google's attempt at a private contractual resolution with a group of publishers and authors in a settlement agreement, with a larger focus on the likelihood of the legality of its actions under copyright law.

In Section II of this Part, I will recommend solutions for the issues relating to the GBS Project through a suggested set of legal reform within the existing Internet intermediaries' law and copyright framework in relation to Google's functions, in a manner that will also be applicable to similar functions offered by other Internet intermediaries. Other legal issues and developments relating to its functions under the project will also be highlighted and some suggestions made thereto for the sake of comprehensiveness, although they will largely be outside the scope and focus of this paper. The proposed reforms will involve either a stronger form of protection through new safe harbor provisions, statutory exemptions or a more comprehensive list of fair use factors.

[1] See, generally, Juliet Dee, *Fingerprints, Grace Notes and YouTube: The Problematic Relationship between Convergence and Copyright Law*, Paper presented at the annual meeting of the NCA 94th Annual Convention, in San Diego, California on 20 November 2008, synopsis available at: http://www.allacademic.com/meta/p259542_index.html. This paper begins with a legal analysis of copyright infringement cases during the past decade, focusing on the influence of convergence and new media such as streaming video. Google Books Search has faced protests from The Authors Guild, and Viacom has sued Google's subsidiary, YouTube, charging copyright infringement. One of the questions considered is whether the Digital Millennium Copyright Act's safe harbor provision requiring notice-and-take-down (the notice process) is adequate in protecting copyright holders' rights. See, also, Pamela Samuelson, *Preliminary Thoughts on Copyright Reform*, Utah L. Rev. 551 (2007), available at: http://privateweb.law.utah.edu/_webfiles/ULRarticles/156/156.pdf and David Kohler, *This Town Ain't Big Enough for the Both of Us - Or Is It? Reflections on Copyright, the First Amendment and Google's Use of Others' Content, Duke Law & Technology Review, No. 5 (2007)*.

3. Section I - the Google Books Search Project (GBS)

The analytical focus of this Part of the paper is on the legislative and judicial response to the GBS under the copyright regime rather than on the private contractual settlement proposed between Google and various groups of copyright owners or on anti-trust issues.[2] However, the provisions of the settlement agreement will be considered insofar as they provide an insight into what the parties consider as fair *vis-à-vis* one another,[3] and anti-competition concerns and the orphaned works dilemma will also be considered incidentally in the course of this analysis. An assessment of the potential hurdle that the Berne Three-Step test may pose to the proposed law reform in this paper will also be made, which will likewise be applicable to the arguments favoring statutory reform in Part One of the paper.

As the search engine that has a dominant position in the marketplace of information, which functions as the main indexer and gateway to WWW content on the Internet, it makes perfect sense for Google to maintain that position by a multi-pronged strategic approach: Constantly develop technologies to improve the accuracy and efficiency of its search algorithms and a user-friendly and individualized interface,[4] create the most comprehensive catalogue and cache of informational archive, diversify its search engines according to the type of work that a user may seek (and in its business portfolio in that regard) and leverage on its strengths as a search engine to appeal to paying advertisers that target the same or similar audience.

With respect to some of its initial, fundamental and basic functions, information locator tools like Google has generally found favor in both policy and law, to the extent that they are the beneficiaries of statutory safe harbor provisions from civil liability for the information that they gather and index for users. But as they diversify and expand in their operations and as their role and objectives change, their functions begin to test the outer limits of these protections and their justification as they conflict with competing public and private interests.

Thus, with regards to the practices of caching and storage, multimedia search engines and keyword advertising has faced legal threats and challenges based mainly on copyright and trademark infringement grounds. We have already seen an example in Part One relating to the launch of Google's images search engine. Google's practice in relation to the caching and storage of literary works has faced different results. Although they may seem alike in operation and general objectives, they differ in relation to the content, the motives and intentions of the parties and in the manner of operation. Before going into the issues relating to the GBS, the concepts of "caching" and "storage" will first have to be defined in the context of this paper.

"Caching" is information uploaded primarily by copyright owners or third parties, which is created and stored in the poster's computer or digital storage devices and subsequently uploaded and stored in the Internet Service Provider (ISP) servers accessible through websites that they create and administer with digital addresses that they purchase and maintain. Google's value-added practice is in creating copies of works in their databases to store for access even after the original post is removed.[5] Hence, any objection will be based on the lack of expressed or implied license to access these works after it is taken down, more so than while the work is still publicly available on the WWW. In *Field v. Google, Inc.*,[6] the US district court for the District of Nevada ruled that Google's caching and display of websites from cache memory was a fair use and hence it was not liable for direct copyright infringement.

"Storing" for the purposes of this Part will has a different meaning and purpose in relation to Google's practices. It relates to the active identification, selection, retrieval, duplication and digital storage of otherwise offline non-digitized documents as a natural prelude to the indexing and display of materials, in

[2] For a critique of the problems in leaving the matter to private settlement, see, Pamela Samuelson, *Legally Speaking: The Dead Souls of the Google Book Search Settlement*, Vol. 52 Iss. 7 Communications of the ACM (July 2009), on why the Google Book Search settlement agreement under consideration could result in an extensive restructuring of the book industry and arguing caution on the private settlement of a public issue or problem, especially with implications on competition and on the rights of orphan book authors that are included in the settlement *in abstentia*. It also shows the problems with leaving what is essentially a matter that should be resolved generally and publicly through laws to private contractual settlement.

[3] For a more supportive analysis, see Matthew Saq, *The Google Book Settlement and the Fair Use Counterfactual*, New York Law School Law Review (2010), available at: http://works.bepress.com/cgi/viewcontent.cgi?article=1006&context%3Dmatthew_sag, where the writer uses fair use to evaluate the settlement terms. The writer analyses the provisions of the settlement agreement and compares it to the likely litigation outcome based on fair use. The settlement provisions are useful as they constitute compromises between the parties based on what they consider fair to both and hence can be used to help determine what will be fair practices that will lift Google's fair use defense, such as the sharing of profits, the creation of institutions to manage the rights and interests of rights holders and the special treatment of orphaned works. Hence, although the writer writes from the perspective of testing the settlement on fair use principles, this article does the opposite, which is to utilize some of the ideas and concepts that render and bolster fair use with recommendations for reform, and proposes practices for Google to adopt in relation to works it intends to include in its GBS project in order to best protect itself and strengthen its position should litigation ensue (irrespective of settlement with a segment of the author/publisher market and of jurisdiction).

[4] Juan Carlos Perez, *Google Revamps Search Results Pages* (PCWorld, 5 May 2010), available at: http://www.pcworld.com/businesscenter/article/195652/google_revamps_search_results_pages.html.

[5] Google has provided its cache since 1998. Google's main search engine scans the web using a "web crawler", an automated search tool which continuously scans available webpages on the Internet and catalogs them into its user searchable web index. In the process, Google also caches the website's source code in a temporary repository on its servers. When Google's search engine runs a text search, links to the cached copy of the websites are returned in the index alongside the full URL for the original page. By clicking on the cached link rather than the original URL, the user views an archived copy of the webpage as it appeared the last time the website was visited and analyzed by the web crawler. Along with the cached page is a disclaimer that the user is not at the original website and hyperlinks that may lead to it. Other major search engines, including Yahoo! and MSN, also uses the same method of caching and featuring webpages.

[6] 412 F. Supp. 2d 1106 (D. Nev. 2006). This case appears to support the opt-out approach. It is for content owners to take either technical measures or make a request to Google following an online process to remove the cache.

whole or in part, online. In 2004, Google began to collect, scan and store literature without prior authorization from their owners. It included content that could fall along the whole spectrum between non-copyrighted works, such as works falling within the public domain, to copyrighted works including orphaned and claimed works. The latter led to various degrees and types of copyright infringement. The difference from its earlier practice of caching is that here, Google actively and with the participation and consent of a select group of libraries, scans printed works without first seeking the expressed permission of individual authors or rights holders or through any possible implied license based on from the digitization of works and the public availability of such works on the WWW.[7]

3.1. *Brief history of development*

Google began as a search indexer and engine for information on the WWW, a large percent of which were literary. It also began Google News, a news aggregator service for users in April 2002.[8] Google Cache, Archive.org and other similar operations, replicates and caches web pages as backups on its servers for access by users after the original website is removed.[9] Meanwhile Google had been updating and improving its search algorithms and display interface constantly to stay ahead of the competition. The GBS that comprises of the Google Books Library Project (GBS) and the Google Partner Program (GPP) was announced and started by Google in 2004.[10] The GBS,[11] which involves the scanning and indexing of book collections, was a project headed by Google with the necessary cooperation of major libraries to include their collections in Google Books by scanning their collections into its database, which "like a card catalog, show users [bibliographic] information about the book, and in many cases, a few snippets - a few sentences to display the search term in context."[12] It displays and allows the download of entire books if it is out of copyright and in the public domain, otherwise there will be links to online retailers or libraries for digital purchases or 'loan' respectively.

As noted, Google Books functions like a library catalogue but in the context of the 'Internet library'. Google invests in and provides the infrastructure and technology to do so, while the library partners obtain the benefits of digitization of their collections in return to further their larger objective to improve access to, and the preservation of, the knowledge contained in these books. After it indexes the books and places them in its database, Google search users can use keyword searches to find words and phrases as they appear within these books. The main view offered of literary works included under the project involves showing only bibliographic information about the book, in particular the title, author, publication and content, and paragraph snippets showing the keywords in context.[13] Currently, entire books will only be displayed and downloadable if and when they are released from copyright protection, automatically under law or voluntarily by the rights holder (i.e. the author or publisher). They will also provide links directing the searcher to online bookstores and libraries for purchase and loan respectively. Presumably, Google is leveraging on this function in anticipation of supporting its profit-driven and diversification goals, by launching Google eBooks (previously Google Editions) that enables Google Inc. itself to sell digitized books, while also earning revenue from its ads placements and library subscriptions. Meanwhile, the official 'goal' touted by Google is to "protect mankind's cultural history".[14] The main benefits are likely to be in relation to the preservation of and access to out-of-print books.

Soon after Google Books was launched, it faced threats of litigation,[15] which has since led to negotiations and a settlement proposal in the US with a group of rights holders,

[7] Contrast this to the opt-in approach of Amazon's "Search Inside the Book" feature and the Open Content Alliance's "Internet Archive" which are essentially based upon an opt-in consent–based model with the cooperation of the rights holders required before scanning. See, Jennifer Suzanne Bresson Bisk, *Book Search Is Beautiful?: An Analysis of Whether Google Book Search Violates International Copyright Law*, 17 Alb. L.J. Sci. & Tech. 271, 275-80 (2007). The writer also analysed the legality of the opt-out model vis-à-vis the International Copyright treaties and found that a mixed system of opt-out for orphaned works and opt-in for copyrighted works with identified authors or publishers (and presumably 'no-opt' for public domain works) may be the best possible solution. *Ibid.* at 305-10.

[8] See Andreas Pouros, *Google's Collaboration in News is Both Good and Inevitable* (Journalism.co.uk, 17 December 2009).

[9] The "Googlebot" which indexes webpages into its central database also saves the HTML portion which is the text and payout without the images.

[10] See the Google Books Search feature at: http://books.google.com/. It was formerly called "Google Book Search" and "Google Print". The sister project is the Google Print Publisher Project (GPP) under which a publisher of a copyrighted book authorizes Google to scan the full text of a published work into the Google searchable database for users to conduct searches that will yield results with information on publications and relevant text as well as a commercial link for purchase. There are no copyright issues implicated by this project as it is opt-in and operates under agreement between the publisher, which has the rights from the author, and Google. See also Kalev Leetaru, *Mass Book Digitization: The Deeper Story of Google Books and the Open Content Alliance* (Vol. 13 No. 10 First Monday, 6 October 2008), available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2101/2037f, for a comparison of the Google Books project to the Open Content Appliance. For an example of access digitization but one that uses an opt-in model, see the Open Content Alliance website at: http://opencontentalliance.org.

[11] See the thumbnailed cover pages of books and how they are displayed on a digital bookshelf at: http://books.google.com/books/.

[12] See "About Google Books" at: http://books.google.com/googlebooks/library.html.

[13] *Ibid.*

[14] See also, archival projects (e.g. the Internet Archive, Microsoft's Live Search Books, Europeana, Gallica, Hathitrust, etc.) mainly for preserving works, and access projects on websites (e.g. Project Gutenberg, etc.) mainly to provide access to out-of-copyright works.

[15] As the project was started in the US, the leading cases also emerged from the US courts. See, *Authors Guild v Google Inc,* United States District Court for the Southern District of New York, Docket No 2005 CV 8136, filed September 20, 2005 and *McGraw-Hill Cos, Inc v Google Inc,* United States District Court for the Southern District of New York, Docket No 2005 CV 08881, filed October 19, 2005. The other plaintiffs include Pearson Education, Inc., Penguin Group (USA) Inc., Simon & Schuster, Inc. and John Wiley & Sons, Inc. See also, Lawrence Jordan, *Communications, Entertainment, and Sports Law: The Google Book Search Project Litigation: "Massive Copyright Infringement" or "Fair Use"?*, 85 MI Bar Jnl. 32 (September 2007). See, further, Kinari Patel, *"Authors v Internet Archives": The Copyright Infringement Battle Over WEB Pages*, 89 J. Pat. & Trademark Off. Soc'y 410 (May 2007) and Branwen Buckley, *SueTube: Web 2.0 and Copyright Infringement*, 31 Colum. J.L. & Arts 235 (Winter, 2008).

although there is ongoing government investigations and resistance from other rights holders and interested parties. The matter is still unresolved at the time of writing this article.[16] Currently, the settlement agreement is before a New York federal district court awaiting approval and it is likely, whatever the outcome, to face an appeal. Many relevant parties from different societal groups with differing views have registered an interest and have spoken out either for or against the project. There are two basic objections to the practice made by its detractors. First, investigations have arisen from the government and complaints have been raised from competitors based on anti-trust or anti-competition concerns, which are currently being looked into especially in relation to the proposed settlement between Google and US authors and publishers.[17] Second, copyright infringement litigation had arisen that led to the negotiations for a settlement, and objectors remain amongst those that reject the settlement outcome, whether on some of the terms or in its entirety. There are also those with other concerns such as privacy and other civic groups.

3.1.1. *The opt-out feature*

Google Books and how it operates presents the clearest example of the opt-out model in effect. Google takes the pre-emptive measure of selecting and copying or scanning books without first seeking permission from copyright holders. Copyright holders can also choose to opt-in through voluntary relinquishment of their otherwise exclusive rights by joining the partner program or by entering into a settlement agreement with Google. Otherwise, they are expected under the current scheme to opt-out of the project by demanding the deletion and removal of their works that are scanned and uploaded without their permission through a complaint to Google.

3.2. *The issues*

To understand and be better able to analyze the strengths and weaknesses of each of their cases, it will be useful to break down the acts that could potentially give rise to a cause of action for copyright infringement.

3.2.1. *First issue: scanning of books for archiving and indexing - likely infringement*

The right implicated here is the right to duplicate a work. This relates to the active scanning and storage of hard copy documents in soft copy form, arguably a necessary action for search optimization and comprehensiveness of online information.[18] The main stated objective for Google is to help archive books, and to make them identifiable and more easily accessible to the general public, through channels such as Google itself and other search engines, library subscription or purchasing options depending on the copyright status of the document concerned.

Mainly, the objections are on the opt-out model of operation that turns the automatic protection rights model under copyright law on its head. Copyright protection extends automatically and by default and as such traditionally and ordinarily requires opt-in arrangements and requests or voluntary relinquishment of rights such as through commercial sale and licensing or non-commercial licenses such as Creative Commons licenses. The GBS scheme automatically divests the exclusive rights under copyright law from their owners by default.[19]

[16] Steve Alexander, *Digital Books Put on Hold* (Star Tribune, 9 May 2010), available at: http://www.startribune.com/business/93147534.html%3Felr=KArks:DCiU1OiP:DiiUiD3aPc:_Yyc:aUU. Google was sued in 2005 by authors, represented by the Authors Guild, and book publishers, represented by the Association of American Publishers. The two sides reached a settlement in 2008, but a federal court never approved it because of concerns raised by government regulators, public interest groups and Google's rivals. An amended settlement reached last November is before the court, though it also faces opposition. For now, scanned library books that remain under copyright cannot be offered in digital form. Google can show snippets of the books *via* its online search engine, but not the whole text. Universities that sent books to Google get the originals back, but the digital versions of copyrighted books cannot even be offered to library patrons unless the settlement is approved. Microsoft has argued that the settlement "would give Google a de facto exclusive license" to orphaned works, changing copyright law without the consent of Congress. Amazon.com alleged the settlement would not only be unfair to authors, publishers and others, but would allow Google to fix prices for the digitized books unless a copyright holder sets a different price. Public Knowledge, a Washington, D.C.-based public interest group, praised the Google Books project for spreading knowledge but said that Google should not get a monopoly on orphaned works. The Electronic Frontier Foundation, based in San Francisco, objected that Google could track people' digital reading habits.

[17] Google's argument is that the contracts it has with the authors and publishers it is settling with are non-exclusive and that the same service can be offered by competitors. However, the anti-trust issues are more complex, see e.g., Randal C. Picker, *The Google Book Search Settlement: A New Orphan-Works Monopoly?*, Vol. 5 No. 3 Journal of Competition Law & Economics 383-409 (2009); Jerry A. Hausman & J. Gregory Sidak, *Google and the Proper Antitrust Scrutiny of Orphan Books*, Vol. 5 No. 3 Journal of Competition Law & Economics 411-438 (2009); James Grimmelmann, *How to Fix the Google Book Search Settlement,* Vol. 12 No. 10 Journal of Internet Law (April 2009); Eric M. Fraser, *Antitrust and the Google Books Settlement: The Problem of Simultaneity*, Stanford Tech. L. R. (10 June 2009); Mark A. Lemley, *An Antitrust Assessment of the Google Book Search Settlement*, Stanford Law School Working Paper Series (8 July 2009); Einer Elhauge, *Why the Google Books Settlement is Procompetitive*, Harvard Law and Economics Discussion Paper No. 464, Harvard Public Law Working Paper No. 09-45 (30 December 2009) and Christopher A. Suarez, *Continued DOJ Oversight of the Google Book Search Settlement: Defending Our Public Values, Protecting Competition*, New York Law School Law Review (2010).

[18] The problem is with the passivity requirement for statutory safe harbor, as "[t]he DMCA was enacted both to preserve copyright enforcement on the Internet and to provide immunity to service providers from copyright infringement liability for "passive," "automatic" actions in which service provider's system engages through a technological process initiated by another without the knowledge of the service provider." *ALS Scan, Inc. v. RemarQ Cmtys., Inc.*, 239 F.3d 619, 625 (4th Cir. 2001) (citing H.R. Rep. No. 105-796, at 72 (1998) (Conf. Rep.) and Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2860 (1989)). See also H.R. Rep. No. 94-1476, at 47 (1976) (discussing that Congress' intent when drafting the DMCA was, *inter alia*, to protect innovation and technology).

[19] See e.g., Mark Hefflinger, *Critics Say Google Book Settlement Violates International Law Treaties*, 10 May 2010, available at: http://www.dmwmedia.com/news/2010/05/10/critics-say-google-book-settlement-violates-intl-law-treaties.

Table 1 – Google's treatment for different sources of books under the GBS and potential legal resolution.

Archive		Index		Others	
Status of books for scanning and storage	Source of book	Full-text search (invisible)	Display (visible)	Downloading, copying and printing	Payment to rights holder
Out-of-copyright (public domain)	Library Project	Yes	Full	Yes	No
In-copyright with publisher agreement	Partner Program	Yes	Full access or 20% preview	No/Limited	Yes (profit sharing?)
In-copyright, no agreement	Library Project	Yes	Snippets only	No/Limited	No (orphaned works?)

3.2.2. *Second issue: public display of excerpts or documents as search results - likely fair use*

The ultimate objective to the practice of copying without permission is to facilitate searches and identification by Internet users of the source of a term, word or phrase, and consequently the literature on a subject matter. In order for such matching of user to material to take place, excerpts will have to be provided in the index search results similar to what is provided for information already available online. The display of search results comes in two stages, at the search results page stage in snippets or limited excerpts for user identification and WWW 'traffic directions'[20]; and a partial or full display of a document as full document or in limited excerpts in the holding page leading from the search results hyperlink, which are created and hosted by Google, not published by a third party on a public website.

This is more of a fair use issue although it is to be noted that its operation *vis-à-vis* the books under the GBS is dependent on the successful implementation of the project. However, in the larger picture, it is source neutral in that it also, and probably still mainly, serves what are already clearly legitimate information and data that are currently made available on the WWW by creators themselves. Information already expressly or impliedly freely shared online include public domain and voluntarily shared books, journal and magazines and other professional and commercial publications, non-commercial publications including government, civil, religious, and other social websites as well as UGC *via* personal websites and blogs (e.g. Blogger.com), social networking sites (e.g. Facebook and MySpace) and applications (e.g. Twitter).

Both the above issues should be treated separately and independently under the fair use analysis. Their objectives overlap but also differ to some extent. For example, the first issue emphasizes on archival and preservation purposes and does not have a real impact on market value either way, while the second issue is mainly determinative based on the fourth fair use factor test.

As an overview, Table 1 shows Google's treatment for different sources of books under the GBS and the potential approach for legal resolution and private settlement.[21] It is to be noted that these practices are still subject to change pursuant to legal resolution under copyright and other laws, and through private settlements.

3.3. *Safe harbor and the DMCA*

None of the existing safe harbor provisions cover the act of active scanning without authorization by Google, which is a form of primary infringement. The safe harbor provisions protect intermediaries from liability for third party material. Hence, the United States' Digital Millennium Copyright Act (DMCA) safe harbor provisions for storage and information locator services are inapplicable to its practice of scanning books for storage.[22] As for the display rights issue, that has already been noted to be a fair use issue, although its legitimacy or otherwise can also stem from the legality of the copying in the first place.

There are already caching fair use exemption provisions under the DMCA. The system caching exemption limits a service provider's liability for keeping unmodified copies of material for a limited time that is made available by another, which a subscriber directs to be transmitted.[23] Similarly, the transitory communications exemption,[24] limits a service provider's liability when it transfers digital information at the request of another and in the course of that transmits, routes, provides connections and makes transient copies of online data. However, they do not cover the practices under the GBS as they are premised on the basis that: First the works are already made available on the WWW by rights holders; and second, these forms of caching relates closely to the technological necessity in order for the Internet as a whole to function efficiently, and does not relate to the role of the Internet intermediary such as an information locator service *per se*.

3.4. *Fair use and the cases*

In relation to literary works in general, there are both examples of non-infringing and fair uses of works in relation to both of the above practices. For example, works, including books that are already freely made available online by the rights holders could be legitimately copied (and cached) and displayed based on the theory of implied license, and non-copyrighted works such as those that fall in the public domain can be freely downloaded and shared.

[20] Compare this to the thumbnail of images and full image display in a new browser window. There may be more similarities in effect if not in format.

[21] Adapted from Eric M. Fraser, *Antitrust and the Google Books Settlement: The Problem of Simultaneity*, Stanford Tech. Law. Rev. No. 4 (September 2010).

[22] 17 U.S.C. §512(c) and (d).

[23] 17 U.S.C. §512(b). There are also user caching fair use exemptions for the downloading and temporary storage of files in the course of WWW navigation by Internet users such as the temporary Internet files directory on the Microsoft Windows computer systems.

[24] 17 U.S.C. §512(a).

There have actually not been any cases on the GBS that have gone through the courts, mainly due to the ongoing negotiations and investigations as well as the judicial examination into the settlement agreement made between Google and groups of publishers in the US (and other jurisdictions like France). However, an analysis based on the basic principles of copyright and fair use and guidance from *Field v. Google, Inc.*, *Kelly v. Arriba* and the *Perfect 10* cases can also be helpful in this analysis.[25]

Google's principal substantive defense in the litigation thus far brought against it on the GBS has in fact been based on the doctrine of fair use. This is because the safe harbor provisions and the purpose-specific statutory exemptions clearly do not cover both of its practices that are in dispute. From an assessment based on the reading and application of the fair use factors, it would appear that its case of fair use is more defensible in relation to the 'visible' *display* of portions of works as part of its search results,[26] rather than its 'invisible' archival practice of *copying* through scanning (the conversion) and the *permanent* storage of books in electronic form (the storage).[27] Any outcome on its settlement negotiations will only have limited effect on Google's legal standing under copyright law without legislative and judicial protection for its practices; otherwise it is at risk of infringement activity, at the very least during the interim period between the inclusion of a work in the scanning (and archival process) and the act of opting out by objectors (and its removal by Google).

It is to be noted, before consideration is made of Google's practices here against the existing statutory exemptions and fair use defense, is that the fair use test is a flexible tool and have actually itself seen evolution to accommodate new realities since its inception in the form of a doctrine of "fair abridgment",[28] which eventually evolved into the fair use concept that recognizes the utility of protecting some legitimate unauthorized uses of what are otherwise exclusive rights. Thus the concept itself is susceptible to evolution.

Under the general fair use exception in the US,[29] "the fair use of a copyrighted work, *including*…for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include: The purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; the nature of the copyrighted work; the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and the effect of the use upon the potential market for or value of the copyrighted work. The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors." There are also specific statutory exemptions, in particular the exemption for libraries, which may be pertinent to this inquiry as well.

What are the types and forms of statutory exemptions and fair use defenses that Google can rely on, and what are the likelihood and the potential impediments to their success? The following evaluation will be in relation to both these questions (unless otherwise indicated):

1. Specific statutory exemption in relation to the first issue on copying – Google can attempt to 'piggy-back' on the library fair use exemption, perhaps identifying or re-inventing itself as an outsourcing agent for cataloguing library books or comparing itself to a library and functioning as such.[30] The spirit and objective of the library exemption certainly can arguably apply to some extent to the archiving and indexing purpose and practices relating to the GBS.[31] But it is highly unlikely to succeed under this provision as it is currently drafted and in its current form of operation. For example, the exemption is limited to only one copy per book or phonorecord, made without purpose of either direct or indirect commercial advantage, by a library or archive.[32] According to Mary Rasenberger, Policy Advisor for Special Programs in the Office of Policy and International Affairs of the Copyright Office in the US, Google could not be seen as a library, archive or as an outsourcing agent.[33] Similarly, it has been argued by others that the Copyright Act's library exemption does not authorize systematic, deliberate reproduction of multiple copies.[34] In any case, the limitations under the exemption do not serve the purpose and operation of the GBS as a whole.

[25] 412 F.Supp. 2d 1106 (D. Nev. 2006); 336 F.3d 811 (9th Cir. 2003) and 508 F.3d 1146 (9th Cir. 2007) respectively.

[26] Which is form or source and substance or content neutral; that is, it applies to literary works in all permutations, not just in relation to book search, for example, to news, blogs, etc.

[27] A major difference to the *temporary* caching and storage of information already on the WWW obtained from automatic trawling of the WWW, even beyond the 'life' of the material from its original source (see e.g., *Field v. Google, Inc.*, 412 F. Supp. 2d 1106 (D. Nev. 2006)).

[28] See *Gyles v. Wilcox*. (1740) 26 ER 489.

[29] 17 U.S.C. §107. Notwithstanding the provisions of §106 and §106A. The doctrine only existed in the US common law until it was incorporated into the Copyright Act of 1876.

[30] See, Hannibal Travis, *Google Book Search and Fair Use: iTunes for Authors, or Napster for Books?*, 61 U. Miami L. Rev. 87, 123-126 (October 2006). However this safe harbor is unlikely to succeed because Google is not and does not function as a library or its agent, and it has profit-driven motives. See, 17 U.S.C. §108(a)(1)–(2) (2006). The library exemption for reproduction is limited to one copy per book or phonorecord made without the purpose of direct or indirect commercial advantage by a library or archive that is either open to the public or to all researchers in a given specialized field. Also, Google is a borrower and not a library and is digitally scanning copies irrespective of the conditions and limitations contained under the exception.

[31] After the first case of copyright infringement case against a library, the US Congress agreed that a library exemption was necessary. See Mary Rasenberger & Chris Weston, *Overview of the Libraries and Archives Exception in the Copyright Act: Background, History, and Meaning* (14 April, 2005), available at http://www.loc.gov/section108/papers.html.

[32] 17 U.S.C. §108(a)(1)–(2) (2006).

[33] Tech Law Journal Daily E-Mail Alert No. 1321, Library of Congress Comments.

[34] Rebecca Tushnet, *My Library: Copyright and the Role of Institutions in a Peer-to-Peer World*, 53 UCLA L. Rev. 977, 1007 n.110 (2006) (citing 17 U.S.C. §108(g) (2000)). See also, *N.Y. Times Co., Inc. v. Tasini*, 533 U.S. 483 (2001). The court in the *Tasini* case, in *dicta*, stated that the Copyright Act's special authorizations for libraries do not cover libraries reproductions of works because the Act only authorizes reproduction of copyrighted works "without any purpose of direct or indirect commercial advantage".

However, comparing the legislation and policy intent behind this exemption, there is good justification for it to be extended by reform to a more general, institution-neutral, archiving exemption,[35] albeit with some necessary conditions and safeguards. Hence, the same or similar arguments made to support the library exemption will also be useful for the proposal for a new exemption for archiving in general, including search engines as an archiving tool that deserves such protection and that provides the same level of societal benefits as libraries, achieving the same aims of knowledge sharing and storage, perhaps with some limitations on its function and purpose.[36] Similarly, the copying exemptions for backing-up of files and other such specific exemptions that serve the same or similar purpose as the GBS can be used to support such an exemption as well. The archiving and indexing function (as facilitated by the scanning and archiving of books) can be another socially valuable tool and valid exemption, as are the case for exemptions for criticism, parody, satire, analyses, and so on. Moreover, the full copying of books is a necessary prerequisite for a more robust indexing function, where the same or similar considerations and arguments from Part One, in support of the copying and modification of images, can also apply.

2. General fair use exemption – Google can attempt to justify its practices of copying, indexing *and* display under the existing fair use doctrine, assisted by the court's recent more generous interpretations of fair use and its factors in the context of, and application to, technological inventions that have an impact or effect on the handling of works.[37] For example, fair use factors and *additional* considerations can be reapplied in this context in a manner that favor the GBS, supported by the following arguments in its defense:
    a. It serves the public goal of copyright law, to benefit society through access to creative works, which should take precedence over private financial interests. It also serves to promote public access to information, including public domain works and to open up access to orphaned works thereby addressing market failures.
    b. Utilizing an interest balancing and balance of convenience analysis, especially in the automated and content neutral information-gathering context; it is easier for a content owner to block or object, whether preemptively by technology or subsequently by notice or request, than for the search engine to seek permission from each and every content owner for permission to include their works in the project, particularly when there will be problems with locating the copyright holder. The same justification can be extended to similar enterprises for search and find functions.
    c. Applying and extending the transformative use doctrine under the first factor,[38] which is a flexible and

[35] See e.g., the European Bureau of Library, Information and Documentation Associations' (EBLIDA) position statement for the European Commission's Google Book US Settlement Agreement information hearing, Brussels, 7 September 2009, available at: http://www.eblida.org/uploads/eblida/10/1252227760.pdf. EBLIDA is an independent umbrella association of national library, information, documentation and archive associations and institutions in Europe. They promote unhindered access to information in the digital age and the role of archives and libraries in achieving this goal.

[36] Meanwhile, holdout copyright owners argue that Google is not the right agent or intermediary (but rather libraries are) to take on this responsibility and project at the detriment of the copyright framework. Siva Vaidhyanathan, *The Googlization of Everything and the Future of Copyright*, Vol. 4 No. 3 UC Davis Law Rev. 1207 (2007), available at: http://lawreview.law.ucdavis.edu/issues/Vol40/Issue3/DavisVol40No3_Vaidhyanathan.pdf.

[37] For fair use analysis under the existing factors to the GBS, see Hannibal Travis, *Google Book Search and Fair Use: iTunes for Authors, or Napster for Books?*, 61 U. Miami L. Rev. 87, 126-139 (October 2006); Kinan H. Romman, *The Google Book Search Library Project: A Market Analysis Approach to Fair Use*, 43 Hous. L. Rev. 807, 828-843 (Summer, 2006) (focus on the practical economic-and market-analysis approach to the factors analysis to argue that fair use should be made out, especially on the strength of the fourth factor); Melanie Costantino, *Fairly Used: Why Google's Book Project Should Prevail Under the Fair Use Defense*, 17 Fordham Intell. Prop. Media & Ent. L.J. 235, 265-276 (Fall, 2006) (with particular focus on the first and fourth factor tests, determining the tool to be sufficiently transformative and does not usurp the market for purchasing, lending and reading books); Emily Anne Proskine, *Google's Technicolor Dreamcoat: A Copyright Analysis of the Google Book Search Library Project*, 21 Berkeley Tech L. J. 213, 223-232 (2006); Manali Shah, *Fair Use and the Google Book Search Project: The Case for Creating Digital Libraries*, 15 CommLaw Conspectus 569 (2007) (the legal barriers to digital libraries and innovations in information dissemination that are of public benefit should be removed based on the fair use analysis and the "good faith test" in the District Court case of Nevada's case of *Field v. Google, Inc.*, 412 F. Supp. 2d (D. Nev. 2006) at 1118-23, which applied the four factors set forth in the statute along with an evaluation of the defendant's good faith in evaluating whether or not their copying constituted fair use); Nari Na, *Testing the Boundaries of Copyright Protection: The Google Books Library Project and the Fair Use Doctrine*, 16 Cornell J. L. and Pub. Pol'y. 417 (Spring, 2007) (argues that by new functionality that will benefit scholarship and research, the Library Project's use of copyrighted works should be considered fair use); Thomas E. Wilhelm, *Google Book Search: Fair Use or Fairly Useful Infringement?*, 33 Rutgers Computer & Tech. L.J. 107 (2006) (analysing and concluding that insofar as statutory exemptions are concerned, while the Library exemption probably will not be available to Google, the fair use exemption probably will); Aundrea Gamble, *Google's Book Search Project: Searching for Fair Use or Infringement*, 9 Tul. J. Tech. & Intell. Prop. 365 (Spring, 2007) (calling it a close call according to how much weight is put on each stakeholder's interest); Elisabeth Hanratty, *Google Library: Beyond Fair Use?*, Duke L. & Tech. Rev. 10 (2005) (pessimistic of success in fair use analysis but urging a negotiated solution or legal reform to preserve and protect technological innovations with societal benefits) and Kyle Lundeen, *Searching for a Defense: The Google Library Litigation and the Fair Use Doctrine*, 75 UMKC L. Rev. 265 (Fall, 2006).

[38] Articulated in *Campbell v. Acuff Rose Music, Inc.*, 510 U.S. 569 (1994). See also Pierre N. Leval, *Campbell v. Acuff-Rose: Justice Souter's Rescue of Fair Use*, 13 Cardozo Arts & Ent. L.J. 19 (1994) and Pierre N. Leval, *Fair Use Rescued*, 44 U.C.L.A. L. Rev. 1449 (1997). A finding of transformative use greatly impacts the overall fair use analysis favorably, but it is not necessary for a finding of fair use. This doctrine have been expanded to take into account and to accommodate technological innovations such as in *Kelly v. Arriba* (for image search engines). It also squares with the treatment of earlier such technologies in the courts as in the *Sony Betamax* case (for time shifting), etc. *Contra*. the doctrine of consumptive use. A consumptive use is one in which defendant's use of the images merely supersedes the object of the originals, instead of adding a further purpose or different character. Whether a use is transformative depends in part on whether it serves the public interest.

additional consideration to determining the "purpose and character" analysis to include public benefits through the use of copyrighted works that transcends the original works' objectives with "something new" and that advances knowledge and the progress of the arts. A search engine certainly can be argued to provide these benefits through the coordinated and greater reach that works can attain, transcending cost, effort and in some cases means.[39] In *Kelly v. Arriba,*[40] the Ninth Circuit case that held the display of thumbnail images in an Internet search engine to be a fair use based on this doctrine.[41] In *Field v. Google, Inc.*,[42] the district court for the District of Nevada also ruled Google's storage and display of websites from cache memory to be fair use. These cases, the reasoning of the parties and the approach the courts have taken in applying the law are pertinent to the issue in this case.[43]

d. In relation to the second factor on the nature of copyrighted works, an analogy can be drawn to the case of *Harper and Row Publishers Inc v. Nation Enters*.[44] The fact that all the scanned works have already been published also adds to the credibility of the application of the fair use doctrine in this case.[45] In a similar vein, the majority of books scanned are fact-based books that may be more relevant for research purposes and this may also weigh favorably for a fair use assessment as well. [46] Also, there are arguments in favor of such uses with respect to orphaned works, where there are economic arguments of market failure and lack of accessibility under the current copyright system,[47] and rare and out-of-print works as well as works that are difficult to obtain within a reasonable time or at a non-exorbitant price.
e. Taking into account the fact that copying is not the goal but the means ('incidental copying') and that the scanning itself although technically infringing does not have a significant negative effect for copyright owners (which can be compared to the caching and temporary files exemptions), and may even have positive effects for them such as providing free marketing and exposure.[48] On the other hand, archiving and preservation of information for future generations is itself of important social interest. There are other value-added functions arising from the combined effects of increasingly comprehensive archiving and indexing such as a more accurate citation count and cross-referencing of literary works.
f. The non-dilution of the works' value and other fair use factors that favor copying for indexing can outweigh

[39] Hal R. Varian, *The Google Library Project* (2006), available at: http://people.ischool.berkeley.edu/%7Ehal/Papers/2006/google-library.pdf, applying the fair use factors test as applied in the *Kelly v. Arriba* case to the design mechanism and objectives of the GBS; and also noting that "copying is incidental" to the project. The main difference between the image search engine or GIS and the GBS is that, in the case of the former, the content owners voluntarily uploaded their images onto the WWW much like caching, unlike in the case of the GBS, where the books are actively scanned and put into the search databases by Google, albeit also to enable search capabilities.

[40] 336 F.3d 811 (9th Cir. 2003).Arriba Soft's thumbnail-generating Internet search engine (copy, modify and adapt) and display of thumbnail images constituted fair use. This more expansive application of the *Campbell v. Acuff Rose* transformative use test in the Ninth Circuit have not as yet spread to the other Circuits. The Second Circuit, for example, have adhered more closely to a narrow application in the cases of *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 608 (2d Cir. 2006) and *Blanch v. Koons*, 467 F.3d 244 (2d Cir. 2006). See, Matt Williams, *Recent Second Circuit Opinions Indicate That Google's Library Project is not Transformative*, 25 Cardozo Arts & Ent. L.J. 303 (2007). In relation to indexing and display, it is possible for Google to draw an analogy with this case since the use of snippets from the original book may be similar to the use of thumbnails in *Kelly v. Arriba and the Perfect* 10 cases. Similarly, in relation to copying and archiving, it can be argued that the replication and thumbnail repository is required for the indexing function and to fulfill essential comprehensive search engine function of all available online content as well to render non-online content digitized and hence recognizable (giving them greater exposure and reputation) and identifiable (matching content to author and other bibliographic information) on the electronic platform. In anticipation of an increasingly fully digitized era of communication and knowledge-sharing, a ore modern, safer and supplemental storage facility and model is also important. The lessons of the huge loss of literature of the past can serve as a warning to a lack of comprehensive archival facilities and operations (although a different problem of the proliferation of inaccurate and useless information can also emerge as a result, which requires a separate set of solution).

[41] However, in *Perfect 10 v. Google, Inc.*, 416 F. Supp. 2d 828 (C.D. Cal. 2006), the district court for the Central District of California held that the GIS, an images search engine very similar to that in *Kelly v. Arriba,* was likely not a fair use based on distinctions in the facts and functions. Google was more commercial and consumptive in its creation and use of thumbnail images for its image search function and Perfect 10's market in thumbnail-sized images on cellphones tipped the analysis towards unfair use. But see, *Perfect 10, Inc v Amazon.com, Inc,* 487 F 3d 701 (CA 9, 2007). In the appeal, this decision was overturned.

[42] 412 F. Supp. 2d 1106 (D. Nev. 2006).

[43] Cameron W. Westin, *Is Kelly Shifting under Google's Feet? New Ninth Circuit Impact on the Google Library Project Litigation*, 2007 Duke L. & Tech. Rev. 2 (2007).

[44] 471 US 539, 564 (1985).

[45] Hannibal Travis, *Google Book Search and Fair Use: iTunes for Authors, or Napster for Books?* University of Miami Law Rev., Vol 61, 601 (2006).

[46] See also, *Williams & Wilkins Co. v. United States* 487 F.2d, 1345 (Ct. Cl. 1973). The law allows copying without permission if it promotes the public access to information rather than the mere ripping off of creative works. In this case, the court found that a medical library did not violate the copyrights of publishers of medical books.

[47] Steven Hetcher, *The Half-Fairness of Google's Plan to Make the World's Collection of Books Searchable*, 13 Mich. Telecomm. Tech. L. Rev. 1, 76 (Fall, 2006), available at: http://www.mttlr.org/volthirteen/hetcher.pdf. ("courts are social welfare maximizers"), arguing that the major category of works owned by publishers should not be a fair use, but that the same transaction cost argument against fair use in the context of the publisher texts would support it in the context of orphaned works (market failure situation).

[48] Kinan H. Romman, *The Google Book Search Library Project: A Market Analysis Approach to Fair Use*, 43 Hous. L. Rev. 807 (Summer, 2006).

the amount and substantiality of the copied work.[49] The same factor is also less unfavorable with respect to public display because of Google's current graduated policy and practices on it. This factor has to be analyzed in the context of other factors. For example, it can also indicate the likelihood and degree of market harm under the fourth factor.[50]

g. In relation to commercialism and market effects, which is the fourth elucidated fair use factor, the effect on the work's value (both commercial, meaning financial market value; and non-commercial, in terms of popularity and exposure) is one of the stronger arguments that can be made by Google based on the extended reach that its platform provides through indexing and Internet access, which are more tangible and provable bases than in the *Sony Betamax* case itself that still yielded fair use as a successful defense. In evaluating the fourth factor, courts often consider two kinds of harm to the potential market of the original work: First, courts consider whether the use in question acts as a direct market substitute for the original work; and second, courts also consider whether potential market harm might exist beyond that of direct substitution, such as in the potential existence of a licensing market. These factors appear to favor the GBS.

h. As the fair use factors listed are non-exhaustive, the "equitable rule of reason" that is built-into the fair use doctrine allows for flexibility in additional reasoning.[51] Indexing is certainly significant and navigability is important as the Internet and the body of materials, with varying degrees of quality, which are available on the WWW increases exponentially. Knowing that there are actually materials available on a subject, identifying sufficient information in order to obtain it, actually being able to get it at a reasonable cost especially for rare, out-of-print and expensive books and having informed choice and options can empower society as a whole.[52] In fact, the "possibility of obtaining the work or adaptation within a reasonable time at an ordinary commercial price" is already a fifth factor in some jurisdictions.[53]

The court's customary favorable slant towards useful technologies that empowers individual users *vis-à-vis* the protection of creative works can work to Google's advantage. This can be seen from the trend of cases since the *Sony Betamax* case, which have been credited for the flourishing of modern communications technology and that have led to subsequent innovations in the field and the IT boom.

3.5. *Other developments and issues*

3.5.1. *Private settlement agreements*

Currently Google has reached a settlement agreement on the class action suit by the Authors Guild and the Association of American Publishers.[54] However its legality is still being tested in the courts. It is awaiting approval in a federal district court in New York and is not resolved at the time of writing of this article.[55]

The proposed settlement arose out of the case *Authors Guild, Inc. et al. v. Google Inc.*,[56] which Google is trying to settle without admitting liability.[57] The settlement is mainly premised on a profit-sharing basis with Google sharing advertising, institutional subscription and book sale revenue with rights holders. The settlement is subject to approval by the courts and also from the governmental anti-competition watchdog. Under the terms of settlement the snippets displayed will remain the lowest common denominator. There are also other heightened forms of display including for preview use (the extent of which is determined by individual rights holders) and public access service (for users at public libraries in accordance with the library fair use). There is some form of choice or control offered to rights holders to claim and "manage" their works, which will be administered by Google and its agent such as determining how much

[49] This has not of itself barred a fair use determination. See e.g. *Sony Corporation of America et al. v. Universal City Studios, Inc. et al.*, 464 U.S. 417 (1984) and, *Kelly v. Arriba Soft*, 336 F.3d 811 (9th Cir. 2003).

[50] *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 587 (1994).

[51] *Stewart v. Abend*, 495 U.S. 207, 237 (1990).

[52] The US Supreme Court has stated that "[c]reative work is to be encouraged and rewarded, but private motivation must ultimately serve the cause of promoting the broad public availability of literature, music, and the other Arts." *Twentieth Century Music Corp. v. Aiken*, 422 U.S. 151, 156 (1975).

[53] See e.g., section 35(2)(e) of the Singapore Copyright Act (Cap. 63).

[54] See, the Google website on the "Google Books Settlement", available at: https://sites.google.com/a/pressatgoogle.com/googlebookssettlement/home. Among others, Google plans to leverage on its efforts with payoffs from the selling of database subscriptions to libraries, selling digitized books through its Google eBooks store and placing advertisements against search result snippets of books. It is also at various stages of negotiation and settlement in other jurisdictions. See e.g., Benedicte Page, *French Deal may Break Deadlock between Google and Publishers* (Guardian.co.uk, 18 November 2010), available at: http://www.guardian.co.uk/books/2010/nov/18/digital-deal-hachette-livre-google and Alex Diaz-Ferguson, *Le Google Strikes Deal in La France: Largest French Publisher to Provide Books to be Scanned by Google* (IP Brief, Washington College of Law, American University, 22 November 2010), available at: http://www.ipbrief.net/2010/11/22/le-google-strikes-deal-in-la-france-largest-french-publisher-to-provide-books-to-be-scanned-by-google/.

[55] For an overview of the sequence of events and the state of play thus far, see Jonathan Band, *The Long and Winding Road to the Google Books Settlement*, 9 J. Marshall Rev. Intell. Prop. L. 226 (2009), setting out all the parties and issues as well as the sequence of events from the launch of the project to the current state of litigation and settlement efforts. See also, Lawrence Jordan, *The Google Book Search Project Litigation: "Massive Copyright Infringement" or "Fair Use"?* 86 MI Bar Jnl. 32 (September 2007), for a short synopsis of the facts and issues surrounding the New York litigation between. See further, Jonathan Band, *The Google Library Project: The Copyright Debate*, (OITP Technology Policy Brief, American Library Association, January 2006), available at: www.policybandwidth.com/doc/googlepaper.pdf and Jonathan Band, *The Google Print Library Project: A Copyright Analysis*, available at: http://www.policybandwidth.com/doc/googleprint.pdf.

[56] Case No. 05 CV 8136 (S.D.N.Y. 2005). Lawsuits have also been launched in other countries like Germany, France and China.

[57] See: http://books.google.com/googlebooks/agreement/ and http://books.google.com/booksrightsholders/.

preview to display, whether advertisements can be displayed on its dedicated webpage, whether to allow sharing of book annotation by users, and so on. The most drastic option is to opt-out of the project. These arrangements, however, do not sufficiently take into account the benefits that Google retains for itself from the treatment of orphaned works or rights holders that do not make a claim for whatever reason, although the Book Rights Registry is supposed to help locate rights holders. Also, the main objection is still Google's default practices in dealing with the literary works and the need for rights holders to 'take back' those rights by opting-out, which under existing copyright law is still a right that resides in the latter by default.[58]

Thus, these settlement arrangements do not resolve the fundamental question of the legality of its operations. But they do offer some idea on what practices can render its dealings under the GBS as fair and for Google to reconstitute its functions under GBS including the making of fair remunerations, such as royalty payments for non-public domain works from revenues earned from ads, subscriptions and sales.[59,60] Also, public domain works are clearly permitted while there are also clear societal benefits of an opt-out procedure for orphaned works. Currently, Google offers three choices to copyright owners, which is to join the GPP or partner program (actual opt-in), the GBS (default opt-in) or neither of the two (actual opt-out). There is much scope for it to adapt its practices and policies to better improve its chances of fair use exemption. Meanwhile a reform of the safe harbor or fair use doctrine can be considered in order to accommodate the GBS.

3.5.2. *Google eBooks*

Meanwhile, Google is seeking to benefit from linking its Books Project to the Google eBooks (previously Google Editions) store,[61] and offers its database of scanned books for library subscriptions. The eBooks store was opened to US customers on 6 December 2010. This is not a problem insofar as copyright law is concerned as there is no real conflict of interest, although there is obviously a convergence of interests and symbiosis, in Google's practices with regard to its Books Search program and its diversification into the sale of books arena. It is relevant where profit-driven motives is a factor and there may also be anti-competition concerns.

3.5.3. *Anti-competition concerns*

The other more contentious parts of the settlement, however, pertain to Google's increased dominion in the area of competition law. For instance, the settlement authorizes Google to sell books[62] in the digital form, to be read online *via* Google's servers. With the Registry acting as the go-between, Copyright owners would then be able to set their own prices, or delegate such pricing decisions to Google.[63] Further, Google will also expand on its current services by offering comprehensive subscription services to libraries, companies, colleges, schools, and other institutions.[64] The settlement, further, permits the creation of a huge research database of every book in the collection, allowing researchers to conduct automated studies that involve computer-based analysis of large numbers of books simultaneously.[65]

Although competition law and issues are complex and outside the scope of this article, and while the anti-trust inquiry into the settlement agreement in the US courts remains outstanding, it is contended that there are ways to achieve its objectives and continue its functions without contravening competition objectives. For example, one possible method to achieving both objectives is to require the cross-carriage sale of background information on books (e.g. author, copyright duration, etc.) and scanned databases such as through a compulsory licensing scheme.

3.5.4. *The interests of creators of "orphaned works"*

Orphaned works is of particular concern as it encompasses a group of creators that may be disenfranchised by the proposed reform. However, legislative and executive solutions can be found.[66] For example, a form of "unclaimed works fiduciary" may perhaps be the solution. Legislation can create an independent institution to delegate the responsibility of the collection of a fixed rate of royalties under a statutory royalty scheme for the use of orphaned other works and

[58] Google still asserts the fair use right in this case to have copied the work in advance before the opt-out is made. This allows Google to act first and refrain later. Under the current settlement provisions, it is stated that "Although Google has no obligation under the settlement to comply with such request, Google has advised the Settlement Administrator that its current policy is to voluntarily honor such requests and refrain from digitizing your books or, if they have already been digitized, refrain from displaying them".

[59] Right to digital publication can be separated from that for the physical analogue and both can be licensed separately or collectively. See, *Random House Inc. v. Rosetta Books* 150 F. Supp. 2d 613 (2001) and *New York Times v. Tasini* 533 U.S. 483 (2001). See further, Robert A. Preskill and Charles McCarthy, *Google Print: Snippets of Infringement*, 23 Ent. & Sports Law. 1 (Summer, 2005), on author's rights *vis-à-vis* publishers and Google in relation to the GPP program.

[60] Even the yellow pages or other indexers have direct or indirect financial motives through their services. The issue here is whether and how copyright owners should be remunerated for the indexing of their works (which requires wholesale copying, even if not display) for which they also obtain a benefit from a potentially bigger market.

[61] Originally slated to operate from July 2010, see Stuart Turton, *Google Becomes Book Seller in July* (PCPro, 5 May 2010), available at: http://www.pcpro.co.uk/news/357712/google-becomes-book-seller-in-july. The e-books will be "device agnostic", in other words it will be fully compatible and device neutral.

[62] Settlement Agreement, *Authors Guild v. Google, Inc.*, No. 1: 2005cv08136 (S.D.N.Y. Oct. 28, 2008), at § 4.2.

[63] *Ibid.* at § 4.2(b).

[64] *Ibid.* at § 4.1.

[65] *Ibid.* at § 7.2(d).

[66] Ryan Andrews, *Contracting Out of the Orphan Works Problem: How the Google Book Search Settlement Serves as a Private Solution to the Orphan Works Problem and Why It Should Matter to Policy Makers*, 19 Cal. Interdis. L.J. 97 (Fall, 2009), suggesting as a viable alternative to a statutory solution the private market solution offered by Google in its Book Search Settlement, which creates a new Collective Rights Organization called the Book Rights Registry to locate and pay rightsholders for Google's use of their works, and which gives Google the right to make productive use of orphaned works without the need to first locate and obtain the permission of the copyright holders of the orphaned works.

licensing these works to third parties for uses that would otherwise constitute infringement, from which authors can then claim royalties upon adequate proof and prospectively re-claim their rights as well (if the rightful owner of so-called "orphaned works" emerges to lay claim to them).

3.6. *Problems with current state of law*

As noted, without any safe harbor protection or specific statutory exemption, the only recourse for Google is to acquit itself under the general fair use exception, which is not likely to protect it from primary infringement liability at least for the first issue on copying. Also, pending a legislative decision on it, Google faces the fundamental problem that is its opt-out model of operation which is in direct conflict with copyright law. It has already faced and deflected one major lawsuit on copyright infringement for the time being but it will likely continue to face increased challenges and criticisms on its GBS as it expands its operations in volume and jurisdictionally.

In relation to the first issue, the private solutions are piecemeal, arbitrary and do not take into account the larger policy and social interests and objectives. They are jurisdiction-based and can develop into a complicated network of disparate agreements with different terms. The most appropriate solution will be statutory amendment of either the safe harbor privileges or a new exemption provision. With regards to the second issue, the institution of fair practices together with more clarity in the law, under the general fair use analysis and an augmented list of factors can resolve the issue and is the more appropriate recourse. This augmented list of fair use factors will complement and overlap with those that will be just as applicable to the GIS, GVS and other search engine operations.

3.7. *Proposed solution under the law*

The interests of society and the consumer are paramount, and private interests are subsidiary and tied to those interests. Even copyright law itself, which is primarily to incentivize creativity by creating property-like interests and rights in creative works, is ultimately with the purpose of benefiting mankind. One of the most important things about the digitization of books and the practice of 'digitally re-mastering' library catalogues is the ease and ability to store in multiple copies (duplicate) and to share (distribute) creative works, and most importantly, to archive and save works that may easily be lost. Universal access to knowledge over time and space is at stake.

The solutions for both issues require different amendments in the law. The first issue regarding copying is better approached as a safe harbor or specific exemption based on archiving objectives while the second issue, which can be developed into fair practices can be resolved through more fair use considerations and changes in Google's practices itself to render it compliant with the fair use factors. Straddling both is the indexing function, which should also be protected by either approach or both. The recommendations are as follows:

1. In relation to the practice of scanning under the GBS, the law should provide a safe harbor provision for the duplication for storage of works for archiving, indexing and cataloguing either as a safe harbor or through a specific statutory exemption,[67] rather than relying on the general fair use analysis and on existing and additional factors to determine the issue. Similar to the proposal for GIS, the same "removal request process" can also operate here to allow copyright owners to opt-out of the project.
2. In relation to the practice of displaying various portions of excerpts or extracts of the books under the GBS, and for information results generally, a reformulation and re-interpretation of the copyright fair use factors is most suitable.[68] There are two approaches, which are complementary and can be applied cumulatively:
   a. The first approach is to update the fair use factors by supplementing the existing factors with additional considerations through expressly incorporating additional factors, and amending the scope and objectives of the existing ones to take into account the benefits of technological innovations and to bring into the equation other important considerations, some of which have already been recognized by the courts. The reform should be through new or augmented factors that will accommodate the functions of search engines such as the display of search results based on the understanding that the ultimate test of fair use and the reason for its interpretative flexibility is that it concerns whether the copyright law's goal of promoting human creativity are better served by allowing the use relative to preventing it. The factors provide guidance that leads to greater certainty and more cohesive decisions. They should remain non-exhaustive.
   b The second approach is the interpretation of the existing fair use factors. This is not really law reform, but more like part of an ongoing process of organic re-interpretation of the flexible fair use doctrine for the Internet context. This has been done most notably in

[67] For a law and economic analysis in favor of fair use, see, Frank Pasquale, *Copyright in an Era of Information Overload: Toward the Privileging of Categorizers*, 60 Vand. L. Rev. 135 (January 2007), available at: http://law.vanderbilt.edu/publications/vanderbilt-law-review/archive/volume-60-number-1-january-2007/download.aspx?id=2534. The writer proposes a revision of fair use doctrine favoring independent categorization. Search engines help sort out the overload of information, the maximizing paradigm and reduction of search costs. "[C]ategorization projects are so necessary to counteract the negative effects of information overload that they deserve positive recognition in the *first* fair use factor". *Ibid.* at 184. On privileged fair use for categorizers (perhaps even to be manifested as a specific fair use provision for categorizers) that can produce 'sub-factors' like transformative use and commerciality are to the first factor analysis (e.g. privilege for categorizers should be proportional to the scope of works categorized and the openness of the categorization project). *Ibid.* at 186-9. See also, Douglas Lichtman, *Copyright as Information Policy, Google Book Search from a Law and Economics Perspective, Innovation Policy and the Economy* (Vol. 9, Josh Lerner, Scott Stern, eds., NBER, 2008), available at: http://www.ipcolloquium.com/Programs/Media/Lichtman_on_GBS.pdf.

[68] See, Hannibal Travis, *Google Book Search and Fair Use: iTunes for Authors, or Napster for Books?*, 61 U. Miami L. Rev. 87, 139-151 (October 2006).

the US Supreme Court *Sony Betamax* case and more recently in the *MGM v. Grokster* case. Google can also justify its practice of scanning as a form of 'intermediate copying for archival purposes',[69] making a distinction between technical and substantive infringement or copying.[70]

There is thus a hierarchy of protection by the proposal of immunity for archiving and indexing functions and fair use defenses for display (i.e. indexing with snippets with no subsequent display or non-substantial follow-on display), which constitutes the minimal threshold of permissible activities based on social importance. Other rights can be left to market forces and private agreements (i.e. substantial follow-on display and full book display for free, subscription or purchase as determined in an anti-competitive and non-monopolistic manner).

3.8. *Private solutions: Google's strategy, legal tactics and best practices*

In the meantime, Google will have to work within the confines of the existing law and it should improve its fairness quotient by business strategy and market them that way as such (which Google is actually already doing) such as by: Offering a platform for the outreach of creative works by creators and access to knowledge by society, which is the very purpose of the copyright law; preserving rare and out-of-print works and making them available and more accessible; offering free access and services to the blind; making available only snippets of copyrighted works; promoting works and linking to retailers (benefiting authors and publishers); other actions that will cast it in a more favorable light under the fair use factors test (e.g. preview, new marketing and e-book alternatives); and generally acting in good faith and showing diligence in producing technology and best practices (e.g. a streamlined complaints process) that can help resolve the issue and reduce the conflict. These acts are also useful and supportive arguments for seeking an overhaul of the copyright regime or the extension of immunity and exemptions. In the process, Google should consider the needs and concerns of both authors (including moral rights) and publishers (profit-driven) that are the main opponents of its project to better alleviate their grievances and minimize threats of legal action against itself.

Google should also segregate and provide differential treatment for different types and categories of works to maximize the benefits of non-infringement (i.e. falling outside coverage of copyright law) and the protections of the fair use doctrine (i.e. falling inside coverage of the fair use defense).[71] For example, it should make available only snippets of copyrighted works, while it can offer greater portions and even full transcripts of public domain works and other works that enjoy lesser protection.

4. Section II - Considerations and recommendations

The important assessment throughout this paper is to determine if Google as an Internet intermediary is a facilitator and hence should be protected or a perpetrator and should be sanctioned. The proposal will constitute a compromise that allows Google to function as it does currently while placing a level of responsibility for it to earn legal exemption from infringement.

4.1. *Fundamental policy considerations*

The fundamental policy considerations have been canvassed in Part One of this paper. To reiterate, both public and private interests should be taken into account in relation to the creation of such technologies and their benefits to society as well as the effects they may have on copyright objectives. Public interest considerations will include those previously stated in Part One. Particularly of relevance here include the storage of knowledge and organization of information. Private interest considerations basically relate to those of concern to the relevant Internet intermediaries and copyright owners.[72]

4.2. *Summary of approaches to law reform*

Insofar as approaches to law reform are concerned, similar to the recommendations in Part One, the most feasible approaches will be either to extend statutory safe harbors, create new exemptions or to refine and expand the fair use factors to provide more certainty and better guidance in determining the legitimacy of a technological function.

There should be a comprehensive and concerted solution and answer to every aspect and performance of a search engine such as those identified in this paper and a 'content neutral' solution will be useful to meet potential disputes in

[69] The amount of copying is necessary to create an efficient directory, the material at this point is not provided to users and is shared with libraries and the copyright owner's market for the work is not negatively impacted but may instead benefit from greater exposure.

[70] Determined to be fair in *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 820-21 (9th Cir. 2003) (deciding that the copying of entire photographs in order to make thumbnails did not weigh for or against Arriba Soft with respect to factor three) and *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1514-16 (9th Cir. 1992) (detailing the process used to copy the code).

[71] Graduated response to display of works (e.g. snippets for copyrighted works to the full display for works in the public domain). Moreover, more studies can be conducted to distinguish the nature and types of works and how they stand up against a fair use analysis, including whether the work is in the nature of a magazine or journal article or a book or treatise, fiction or non-fiction, and so on.

[72] E.g., Jennifer Valentino-DeVries, *Facebook, Yahoo, Other Web Giants Back YouTube in Viacom Case* (The Wall Street Journal, 31 May 2010), available at: http://blogs.wsj.com/digits/2010/05/27/facebook-yahoo-other-web-giants-back-youtube-in-viacom-case/ and Matthew Lasar, *Internet Democracy at Stake in Google/Viacom Lawsuit?* (ars technical, 29 May 2010), available at: http://arstechnica.com/web/news/2010/05/post-5.ars.

the future as technologies continue to emerge in relation to search engine functions. Statutory reform will also be more useful to resolve the issue and avoid uncertainty and unnecessary litigation, which can impede technological progress.

There should also be global consistency and harmonization of policy and laws because search engines like Google operates worldwide, the Internet and WWW functions across borders and transactions are done within different jurisdictions as well as among different parties. This can be done by establishing a worldwide working group or authoritative body under the auspices of an international organization to set rules and guidelines to establish the permissible parameters as well as to determine the legitimate procedures and processes relating to the operation of technologies. The recommended law reform and amendments proposed in this paper should also be taken up by the World Trade Organization (WTO) or the (World Intellectual Property Organization (WIPO) with a view to their incorporation into the international copyright regime.

4.3. *The proposed approach to law reform*

Traditionally, safe harbor extends to Internet intermediaries with respect to their role in the dissemination and distribution of third party material. Their role have evolved within the function of an information locator tool (e.g. for images search) as well as in its diversification of roles that are related to its basic search function (e.g. caching, archival and hosting facilities). These functions can be controversial but are no less important, integral and useful to navigating the WWW. Perhaps it is time that this should be reflected in the safe harbor laws.

A reform of the safe harbor provisions as recommended can be considered a significant move as it will be the first time that such statutory protections covering the copyright regime goes beyond the litmus test of "control" that is the mainstay and the key determinant of protection under current provisions in both the US and EU.[73] It will also be the first foray into the statutory protection of an Internet intermediary for first party content or as the primary party; for example, the actual exercise of certain exclusive rights by Google and other search engines performing the same role or function without first seeking permission from the copyright owners as in the case of the GIS and GBS. Moreover, the conditions for safe harbor relating to "control" such as adherence to the notice process and showing a lack of the requisite "actual knowledge" or "awareness" and "no direct financial gain" will have to make way for more liberal qualifications or conditions such as an objection or 'opt-out' procedure. The effect of such a safe harbor will also entirely protect the search engines as a statutory *right*, which can be used as a *sword* against aggrieved plaintiff copyright owners who would carry the burden of proof to surmount the claim of immunity. It restrains an action for copyright infringement from proceeding if applicable.

Alternatively, the same exemptions can be made within the copyright statute as new and additional specific statutory exceptions or be relegated to a fair use defense. This will provide the search engine and other intermediaries providing the same or similar services the opportunity to *shield* or *defend* itself from a cause of action initiated by the copyright owner. The plaintiff copyright owner has the right to sue and the provisions do not, unlike the previous case, protect the defendant from an action being instituted against it. In this case, the exemption falls within the copyright regime and the defendant bears the burden of proving eligibility to the defense against an allegation of infringement. In this situation, the threat or possibility of a civil suit can have chilling effects on practices, especially amongst less powerful intermediaries. Statutory exemptions are not new and there are many such limitations within the copyright itself as well as in other statutes.[74]

4.4. *New safe harbor provisions or specific fair use exemptions for search engine archiving function and indexing tools*

Together with Part One, I have identified two main functions with significant social benefits that outweigh the protectionist interests of copyright holders and for which statutory general protections or specific exemptions should be extended. They are the archiving and indexing-related functions.

The proposed new statutory provision to exempt archiving-related functions should be subject neutral, thereby extending protection or exemption to any Internet intermediary that performs the same function within the definition and scope of the provision. Thus, the legalization of the archiving function can cover a range of activities from caching practices such as that in *Field v. Google, Inc.*[75] to the practice of scanning and storage under the GBS project and YouTube's practice of providing storage facilities for UGC. This proposed provision will not extend to the display or broadcasting of the content or material, but only to the reproduction and storage of such material.

Similar to the indexing-related functions of its other forms of searches, the index and search functions necessitate the

[73] Currently, the liability of intermediary service providers provisions in the EU E-Commerce Directive (Directive, 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market) and the copyright provisions in the United States' Digital Millennium Copyright Act (7 U.S.C. §§ 512, as amended by the Online Copyright Infringement Liability Limitation Act (OCILLA)) contain conditional safe harbor provisions for basic functions relating to Internet access and indexing including transitory network communications, system caching and basic information location tools (i.e. hyperlinking to illegal third party websites).

[74] Currently, copyright statutes provides limited purpose-specific exemptions such as for criticism or review, for the reporting of current events, for judicial proceedings or professional advice and for collections of works for the use of educational institutions and libraries. More recently there have also been more of such limitations in relation to electronic and digital interests like the defences for making back-up copies of computer programs, decompilation, the use of copyrighted computer programs for observation, study and testing, and so on. Limitations also sometimes exist in other statutes. For example, the United States' Audio Home Recording Act (Pub. L. 102-563, 106 Stat. 4237 (1992)) permits the making of copies of audio recordings for non-commercial personal use in certain circumstances. See 17 U.S.C. §1001-10.

[75] 412 F. Supp. 2d 1106 (D. Nev. 2006).

display of portions of literary works and other relevant information such as bibliographic data. Thus, the safe harbor or statutory exemption for indexing, proposed under Part One for images and related search engines, is likewise applicable to, and useful for, literary and purely information-based search indexing. A simple example of how such a protection or exemption can look is provided in Table 2.

These amendments, together with the existing safe harbor provisions, will effectively permit and legitimize the following activities or functions currently practiced by Google (and the same or similar functions of other online entities) in relation to digital archives cum search engines –

1. Scanning and indexing of materials for the purpose of archiving and user-catalogue identification services with reasonably adequate security (to prevent unauthorized access through hacking and perhaps requiring insurance) (first party duplication and limited display).
2. Providing hyperlinks (including deep links)[76] to third party websites that may contain illegal copies or commercial or non-commercial websites that may offer legal copies of the work (e.g. on request, subscription or sale basis). This exemption applies beyond books search to all literary materials like news agents and blogs (third party material).

4.4.1. *Instituting a "removal request process" in addition to the notice process for protected activities not relating to third party material*

For the same reasons as stated in Part One, the "removal request process" should be provided as a fair compromise, which will allow objectors to opt-out of any aspect of Google's search engine processes through a simple, clear and more formal streamlined process. The provision of such a facility and adequate response and enforcement should be made a pre-condition for eligibility to statutory protection (i.e. safe harbor or specific exemption under the Copyright Act), and is included in the proposed provision in Table 2.

4.5. *Updated fair use factors under the general air use exception provision*

As in the case of Part One, the other option will be to make use of the general fair use defense to extend protections for archiving, and particularly for indexing functions to search engines. The recommended factors will remain the same and apply just as well to the GBS as it does to the GIS. The reasons for the recommended amendments to the fair use provision have already been canvassed in Part One and reference can be made to it.

The proposed amendment to the current fair use provision is reproduced below in Table 3.

4.5.1. *Reconciling the fair use doctrine and its expansive effects with the international obligations under the Berne Three-Step test*

The main intellectual property and copyright-related international law instruments contain a three-step test for assessing the validity of exceptions to the exclusive rights of copyright owners. The same requirements can be found in Article 9(2) of the Berne Convention,[77] Article 13 of the TRIPS Agreement,[78] Article 10 of the WIPO Copyright Treaty,[79] and Article 16 of the WIPO Performances and Phonograms Treaty.[80]

The Berne Three–Step is arguably not an impediment to the creation of specific statutory exemptions in whatever form, whether as a safe harbor protection or as a purpose-specific exemption, provided that the three conditions are met.

First, under these conditions and according to the "fixed view", exceptions should be allowed for specific purposes and not generalized exceptions. Second, the exception must not conflict with normal exploitation of the work, that is, it must not interfere with the normal use of the work by the rights holder.[81] What constitutes normal exploitation can evolve "as a result of technological conditions or changing consumer preferences."[82] Third, it must not be unreasonably prejudicial to right holders' legitimate interests (mainly to do with a significant loss of revenue from the use by the exempted party).

The flexible but general fair use doctrine itself has been the subject of criticism based on the test although it is to be noted that the doctrine has survived and thrived for decades and it has even expanded jurisdictionally and been transposed into the copyright laws of other countries. It is also to be noted that the US was not obliged to amend her fair use provision when she joined the Berne Convention in 1989.

There is a schism between those that prefer a strict and narrow reading of the three step test and who also require the three elements to be individually satisfied and those that have a more liberal and flexible reading of the test that

[76] This excludes 'system hacking' or other technological methods that circumvents limited or secured access technologies that are instituted on certain websites (e.g. requiring subscription and password, acceptance of terms in a click-wrap link, etc.). However, it includes deep linking into websites where no such technology is instituted but that the website creator or administrator intended, expressly or otherwise, for visitors to enter through the homepage or otherwise fulfil other conditions (such as reading the terms of use under a browse-wrap link) for further access. This balances the *convenience and interests* of the content host and service provider *vis-à-vis* the copyright holder.

[77] Berne Convention for the Protection of Literary and Artistic Work of September 9, 1886, July 24, 1971, 1161 U.N.T.S. 3, 31.

[78] Agreement on Trade-Related Aspects of Intellectual Property Rights, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, art. 13, Apr. 15, 1994, 1869 U.N.T.S. 299, 305. This Article was derived from the Berne Convention.

[79] World Intellectual Property Organization Copyright Treaty, art. 10, Dec. 20, 1996, 36 I.L.M. 65, 71.

[80] World Intellectual Property Organization: Performance and Phonograms Treaty, art. 16, Dec. 20, 1996, 36 I.L.M. 76, 85-86.

[81] See, Jo Oliver, *Copyright in the WTO: The Panel Decision on the Three-Step Test*, 25 Colum. J.L. & Arts 119 (2002). "Empirical approach" or "normative approach" or both? *Ibid*. at 155-57. See also, Annette Kur, *Of Oceans, Islands, and Inland Water – How Much Room for Exceptions and Limitations Under the Three-Step Test?*, 8 Rich. J. Global L. & Bus. 287 (2009), for a more flexible interpretation of the test.

[82] *Ibid*. at 158. Existing conditions and usual use or a more forward-looking form of assessment under a "dynamic view"? *Ibid*. at 159.

Table 2 – Template for a safe harbor or statutory exemption provision for the archiving and indexing functions.

Archiving of information[a,b]

A service provider shall not be liable for the copying of any information for the purpose of archiving or for the purpose of making more efficient the function of indexing the information, on condition that the provider, upon receiving a removal request, acts expeditiously to remove the copy of the information,[c] and on condition that reasonably adequate security is put in place to prevent unauthorized third party access to the information.[d]

Indexing of information

A service provider shall not be liable for the modification and display of information in an directory, catalogue or index for the purpose of organizing information, and shall also not be liable for providing the technical means to transport a user to its source, whether or not as an information locator tool in relation to search results, for the purpose of making more efficient the function of indexing the information, on condition that the provider, upon receiving a removal request, acts expeditiously to remove the copy of the information.[e]

a "Information" will include any form or type of online material.

b Although the anti-competition concerns, which is the other major legal issue relating to the project is beyond the scope of this paper, these concerns can be also be separately resolved as noted earlier in the paper such as through the institution of statutory cross-carriage of goods with a royalty or compensation scheme for the mandatory sharing of archived materials with other search indexers as well as for any potential social good or service.

c I.e., request from the copyright owner preferring not to be included in the archive or index, or informing that the source of the information is illegal. There also needs to be put in place a proper set of procedure and process for the filing of such requests (and the evidence or information needed, perhaps in a template) and the response to such a request (by the service provider, including the reasonable reaction time).

d What would legally constitute reasonably adequate security may also have to be statutorily set out in the same manner as the conditions for secured electronic signatures have been set out in the UNCITRAL Model Law on Electronic Signature and in national legislation on e-signatures.

e However, other existing conditions for eligibility will have to be excluded for obvious reasons such as the requirement of an effective policy to deal with instances of repeat infringement (by another) and the support of standard technical measures for the protection of works (except insofar as Google itself does not circumvent such measures). See, 17 U.S.C. §512(i).

Table 3 – Example of proposed amendments to the U.S. fair use provision.

The fair use of a copyrighted work is not an infringement of a copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include [*but are not limited to*]:

1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes *and whether such use is of a transformative nature or is for a consumptive purpose*;
2. the nature of the copyrighted work *and the nature of the use, including personal use*;
3. the amount and substantiality
   a. of the portion, *including its* used in relation to the copyrighted work as a whole and substantiality includes format, size, resolution and other features of use that distinguishes it from the original work;
   b. of infringing uses as compared to non-infringing [in relation to the services of a technological medium or media for communication and transfer];
4. the effect of the use upon the potential market for or value of the copyrighted work*; and*
5. the copyrighted work was used in good faith and reasonable and proportionate measures have been taken to protect the copyrighted work from infringement.

recommend a holistic application of the test. Depending on which school of thought, it is argued that the recommendations that are made in Parts One and Two still satisfies the requirements, particularly the suggested statutory amendments to incorporate purpose-specific protections or exemptions. Certainly, it can be argued that both the archiving and indexing functions are special cases and not the usual form of exploitation of the works – in fact these arguably 'transformative' uses are supplementary to the normal uses of the works and are even complementary to and enhances the value and normal exploitation of the works so as to overall promote the legitimate interests (especially commercial interests) of the copyright holder (provided of course that safeguards against potential abuses are instituted to prevent or minimize potential detriment, such as the 'leaking' of such works to the public domain or direct commercial and competitive exploitation of the works in a similar manner as rights holders'). Surely whatever form it takes for the protection of these functions, these legitimate purposes should not be overlooked.

4.6. *Practices determined to be unjustified and not eligible for protection*

Although the suggested reforms made above are favorable to Google there are some current practices and actions that should not be eligible for protection, and hence the industry-wide practices or abuses relating to these practices should cease, unless private arrangements or settlement agreements are made that does not go against public policy interests. For the GBS, it includes providing more than mere snippets of existing copyrighted materials that fail the normal application of the fair use factors test. It may still fulfill the fair use criteria upon the application of the fair use factors on a case-by-case basis if it is operationally optimized and infringement is minimized.

5. General conclusion: a truce between Google and the creative empire with the law as the arbiter – instituting legal support for a fair and just compromise

Search engines such as Google will continue to innovate, and although the ulterior or secondary motives may be financially driven and they may sometimes go beyond what is justified, that does not mean that the technological changes in service do not have great social benefits. Each function should be looked at with a view to policy accommodation or otherwise and these should be clearly enunciated under the law so that they do not stunt the growth of the industry and the advancements in technology, while helping to avoid unnecessary litigation. This should be done periodically and with some frequency – and it is submitted that the safe harbor and fair use provisions *vis-à-vis* search engine technology is overdue for review and amendments given the developments on several fronts as apparent from the above analysis as well as the evaluation of the GIS in Part One.

Before I conclude, some provisos have to be made to the paper as a whole. Although Google is used as the obvious subject for a case study on the policy and law on information locator services and the functions it entails and that are developed pursuant thereto, the analyses and proposals are meant to apply to all search engines and other entities that perform the same functions with the same or similar objectives. Also, although the analysis here is in relation to the US law, with its rich history of litigation and jurisprudential analysis and law, both statutory and judge-made, to regulate and facilitate technological progress and the Internet and WWW,[83] the same issues and problems faced in other countries and regions are largely similar and require a similar consistent.[84] In fact, a harmonized solution on a global scale will do a lot more to resolve the problem, particular since Internet interactions and transactions are essentially worldwide.

Finally, in answer to the question posed in the title at the very beginning of this paper: Google is a perpetrator or a facilitator *as defined by law*. In most cases, Google is a facilitator whose functions should be protected, and where it perpetrates or enables infringement, the protections will not exist and it will continue to face infringement liability under copyright law. The legitimacy of its role and functions can change through law reform. The exercise in this paper is to set out clearly the existing and the proposed boundaries between these definitions as applied to Google's operations canvassed under both Parts One and Two. It is hoped that the suite of recommendations set out in both Parts will be considered for legal reform, or at the very least will serve as a springboard for discussions with a view to legal reform. Lack of clarity in the law, will only lead to more uncertainty and constant disputes over its roles and functions. The analysis in this paper is to find a suitable and fair compromise for all parties concerned that can bring a state of convergence between the private disputants, which will at the same time fulfill the overall objective of law and policy in the case of the copyright regime, which is to optimize the benefits of creative works to society as a whole. Promoting the reach and availability of such works is paramount.

Postscript

On 22 March 2011, the US Federal Judge in New York before which the legal settlement was brought rejected the deal citing copyright, antitrust and other concerns.[85] In particular, the over-extensive effects of the 'opt-out' mechanism, the potential monopolistic effects of the agreement and the treatment of orphaned works and unclaimed books under it were highlighted as causes for concern.[86]

Even though the decision is a substantial roadblock to the continued operation of the GBS, it does not necessarily terminate the future potential for the digitization of literary works. The Judge himself acknowledged that "the creation of a universal digital library would benefit many," and of its societal benefits even when he found that the current terms of the proposed agreement was "not fair, adequate and reasonable." The implication is that an agreement that is fair, adequate and reasonable in the form of a revised agreement could be acceptable. The door is not closed as the judge denied the motion for approval without prejudice, which means that it will reconsider the settlement if the parties can renegotiate a revised agreement. This merely brings the parties back to the negotiating table to work out a more acceptable solution in the eyes of the court. Meanwhile, the other alternative, which was proposed by many opponents of the deal, is for the enactment of legislation to deal with the problems and to offer a more comprehensive, balanced and public solution. The Judge likewise stated that in the matter of unclaimed books that it is "a matter more suited for Congress".[87] This is consistent with the proposal in this paper on just such a legislative solution in the context of copyright law (although other ancillary and additional legislative amendments will also have to be prescribed taking into account all the other issues and concerns).

Warren B. Chik (*warrenchik@smu.edu.sg*) *Asst Prof of Law, Singapore Management University, Singapore.*

[83] Brandon Brown, *Fortifying the Safe Harbors: Reevaluating the DMCA in a Web 2.0 World*, 23 Berkeley Tech. L.J. 437 (2008).

[84] Joris van Hoboken, Legal Space for Innovative Ordering: On the Need to Update Selection Intermediary Liability in the EU, Int'l J. Comm. L. & Pol'y 49 (Winter, 2009).

[85] The full opinion is available at: http://www.nysd.uscourts.gov/cases/show.php?db=special&id=115.

[86] *Ibid.* at pp.24, 28, 31 and 36.

[87] *Ibid.* at p.24. The Judge also stated that the settlement infringes on the legislature's power to address copyright issues. *Ibid. at* p.30.