

By Eleonora Rosati, 21 November 2013

A closer look at the Google Books decision



Last Thursday Judge Denny Chin of the US District Court for the Southern District of New York **[yep, the same one who [rejected the proposed settlement agreement in 2011, holding that it was too unbalanced in favour of Google\]](#)** issued his much-awaited [ruling](#) in the [Google Books Library Project](#) saga, which started back in 2005. It was then that the [Authors' Guild](#) and the Association of American Publishers ([AAP](#)) sued Google for copyright infringement over non-authorized scanning of quite a few books.

Just to avoid any confusion among readers, Google Books is the broader project that includes the Library Project and the [Partner Program](#) (formerly known as Google Print). What is sometimes known as the *Google Books* case just involves the Library Project.

Background

Since 2004 Google has scanned more than 20 million books in their entirety **[with approximately 93% of books being non-fiction, and the great majority of works being out-of-commerce]**, and delivered digital copies to participating libraries **[the New York Public Library, the Library of Congress, and a number of university libraries can download a digital copy of each book scanned from their collections, but not copies from other libraries' collections]**, created an electronic database of books, and made text available for online searching through the use of snippets **[users can search the full text of all the books in the corpus, although it is not possible to view a complete copy of a snippet-view book]**. Some libraries have agreed to allow Google to scan only public domain works, but others have also permitted the scanning of in-copyright content. Overall, libraries have agreed to abide by the copyright laws with respect to the copies they make.

The AAP and Google concluded a settlement [agreement](#) last year ([here](#) and [here](#)), but this did not affect the still ongoing litigation between the Authors' Guild and Google. In particular, the main question left on the table was whether Google could successfully sustain that its Library Project activities were protected as fair use under [§107](#) of the US Copyright Act.

Under US law, the following factors must be considered in order to determine whether the use made of a copyright-protected work may be considered fair:

- the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes **[a key consideration here is whether the use is transformative, ie whether the new work merely supersedes or supplants the original creation or whether, instead, it adds something new, with a further purpose or different character. For recent, yet controversial,**

cases *see* [here](#) and [here](#)];

- 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.

Why Google Library Project is great

Before considering the four fair use factors, Judge Chin highlighted the benefits of the Library Project, including:

- Providing a new and efficient way for readers and researchers to find books (*“Google Books has become an essential research tool, as it helps librarians identify and find research sources, it makes the process of interlibrary lending more efficient, and it facilitates finding and checking citations ... Google Books has become such an important tool for researchers and librarians that it has been integrated into the educational system — it is taught as part of the information literacy curriculum to students at all levels.”*)
- Promoting a type of research referred to as “data mining” or “text mining” **[does this ring any bell to European readers?]**
- Expanding access to books, *eg* by providing *“print-disabled individuals with the potential to search for books and read them in a format that is compatible with text enlargement software, text-to-speech screen access software, and Braille devices.”*
- Helping to preserve books and give them new life, *eg* in the case of out-of-commerce works.
- Helping authors and publishers (*“When a user clicks on a search result and is directed to an “About the Book” page, the page will offer links to sellers of the book and/or libraries listing the book as part of their collections ... Google Books will generate new audiences and create new sources of income”*) **[basically, this means that plaintiffs litigated this case for years against their own interests ...]**

Fair use factors

(1) Purpose and character of the use - *“Google’s use of the copyrighted works is highly transformative”,* the Judge found, in that *“Google digitizes books and transforms expressive text into a comprehensive word index that helps readers, scholars, researchers, and others find books.”* Furthermore, *“[t]he use of book text to facilitate search through the display of snippets is transformative”.* In addition, *“Google Books is also transformative in the sense that it has transformed book text into data for purposes of substantive research, including data mining and text mining in new areas”.* Finally, *“Google Books does not supersede or supplant books because it is not a tool to be used to read books.”* Although *“Google is a for-profit entity and Google Books is largely a commercial enterprise ... even assuming Google’s principal motivation is profit, the fact is that Google serves several important educational purposes.”*

(2) Nature of copyrighted works - Two considerations favoured a finding of fair use in respect of this factor: (1) most scanned works were non-fiction books, and (2) the books were published works.

(3) Amount and substantiality of portion used - Although Google limited the amount of text

displayed in response to a search, the fact that Google scanned full-texts and offered full-text search of books was found to weigh slightly against a finding of fair use.

(4) Effect of use upon potential market or value - Google does not sell its scans and the scans do not replace the books. Above all, *“a reasonable factfinder could only find that Google Books enhances the sales of books to the benefit of copyright holders ... Google Books provides a way for authors’ works to become noticed, much like traditional in-store book displays.”*

Conclusion

Considering all the factors above, the judge concluded that *“Google Books provides significant public benefits”* and granted Google’s motion for summary judgment.

While it is not said that this is the end to the Google Books saga, last Thursday’s ruling certainly represents an important victory for Google.

Looking at the decision through European lenses, two sudden questions arise:

- Are orphan works a fake problem? At least under US law it would seem so, as there is no mention of them in the decision of Judge Chin.
- Are text and data mining activities something which falls outwith the scope of copyright protection *tout court*? From what Judge Chin wrote, it would seem so: text and data mining would require neither a licence nor a specific exception. Although US open-ended fair use clause differs from InfoSoc Directive’s exhaustive list of exceptions and limitations, Judge Chin did not seem to consider that such activities could infringe exclusive rights of copyright owners. From an EU perspective, if one transferred the interpretation of “commercial” provided in the ruling to this context, it could be argued that most text and data mining activities would be already covered by Article 5(3)(a) of the [InfoSoc Directive](#).

[This blog post was originally published on [The IPKat](#) on 17 November 2013]