

Google and Copyright Infringement

Google, back in 2004, began an endeavor to index the contents of an enormous number of books through its search engine, so that google users would be able to full text search books that were otherwise unpublished on the internet. Under U.S. Copyright law, books that were published before the 1920's (and certain texts published after that time that did not comply with the renewal requirements and were not saved by the Copyright Act of 1976) are in the public domain and can be freely copied without the need of prior consent or the paying of royalties to the author or his/her estate. Hence, you can [find a copy of Edward Gibbon's Decline and Fall of the Roman Empire on google's book search](#), because Mr. Gibbon originally wrote the manuscript well before the earliest date that the book could be protected by current U.S. Copyright law. Of course, what got google into trouble was not long dead authors but very alive ones (or ones whose estate or a third party owned a valid copyright to the work), which led to a lawsuit against google in federal court in 2005 by several named plaintiffs and an association, the Author's Guild, who represents over 8,000 other authors. *The Author's Guild, et. al. v. Google, Inc.*, 05 CV 8136 (S.D.N.Y. Sep. 20, 2005).

The complaint in 2005 alleged that google's indexing of these books without paying a license fee to the individual authors with valid copyrights was copyright infringement writ large, and that the indexing was done in search of advertising revenue (an expressly commercial purpose).

Infringing the valid copyright of another without paying the customary license fee is the sine qua non of an unfair use, and I suspect that were we to see this tested in a court, google would likely have lost the suit. However, the case didn't get very far as the parties entered into negotiations to settle the matter. In 2008, a proposed settlement was announced ([see a CNET article here](#)) which would have had google pay the Author's Guild about \$125 million in royalties for google to continue its "exploitation" of works that were probably protected by copyright. The settlement also provided a method for dissatisfied authors to opt-out of the indexing process (which was google's position from the start on this project).

One of my readers commented that google is also providing a service to authors whose works are indexed which is that the search engine makes their works more available to a worldwide audience, and one would think that most authors would appreciate the free exposure by google. In the writing business, exposure usually costs money, and the upfront cost of that exposure is usually borne by the publisher who decides to publish the author's work. In this case, google's advertisers are the ones that are paying for the cost of the exposure. In addition, google's main source for materials was libraries across the country who had copies of books they were willing to share with the google indexing process. An unknown or little-known author might only have his work in one or two libraries, where the google "library" gives access to people everywhere. This is not to say that because google was giving away worldwide exposure at the expense of its advertisers that google therefore had a fair use defense. But, one could argue that authors so indexed were being given some value (in the form of exposure) which might very well otherwise exceed the value of the royalties due on the regular sale of their books, especially if a google user ended up buying a paper copy of another book by the author as a result of a google search. (And, by the way, when I was searching for the [Decline and Fall of the Roman Empire](#), Amazon was

one of the advertisers on the google search, advertising where I could buy the paper copy of that very title).

The opt out provision also allowed authors to leave google's index who could afford to control their exposure, or whose book sale royalties exceed the value of google's worldwide search. The dilemma with this approach is that, if google was found to be infringing the copyrights of others through its creation of complete digital copies of the protected works, the fact that it was providing "free" worldwide exposure would not result in a fair use defense. Under section 107, there are four factors weighed by a court: (1) the character of the use, (2) the nature of the protected work, (3) the amount of the work used, and (4) the effect of the use on the market for the copyrighted work. Given that google was making money on advertising and is a profit-making entity, the first factor weighs against them. A court may take into consideration the value of the worldwide exposure google was giving to the protected works in this factor. Highly creative works would get more protection under the second factor, so this might go either way depending on the plaintiff and his work. But the third and fourth factors would probably weigh against google because it took the entire work and made a digital derivative and didn't pay the customary price in royalties to the author. *See Brown v. McCormick*, 23 F. Supp. 2d 594, 607 (D. Md. 1998).

Besides the problems with its fair use defense, critics of the settlement have expressed concern that google's book collection looks a tad monopolistic (including the Department of Justice, who opened an antitrust investigation according to [Reuters](#)). There is concern in the online community that google may have control of too much information which may ultimately stifle innovation by others.

Monopolizing a market generally violates the Sherman Anti-Trust Act, which can lead the Department of Justice to file suit against the alleged monopolizer. Such suits have caused large companies like IBM and AT&T to either stop seeming to be monopolies, or to breakup outright into smaller units. Last year, anti-trust concerns [stopped google from establishing a search marketing](#) relationship with yahoo, even though google was probably not trying to control the world of search but just trying to help yahoo fend off a purchase by Microsoft (which ultimately did fail and subsequently led to the ouster of Yahoo's CEO and a founder Jerry Yang later in 2008).

Ironically, holders of a valid copyright exercise a legalized monopoly over the thing copyrighted, which, while limited to the duration of the author's life plus 70 years, is a relatively long time. For highly valued items, such a monopoly could effectively stifle innovation, at least for those that wish to make derivative works from the copyrighted work but cannot afford to pay the "customary fee" to the copyright holder. Effectively, the copyright holders represented by the Author's Guild are one set of monopolists fighting with another alleged monopolist, google, which is probably far larger, but probably not otherwise more or less sympathetic. On the other hand, the copyright monopoly does have limits built in to the rights granted under the Copyright Act itself, including fair use under section 107, which provides for some academic and non-profit expression by individuals who would otherwise be copyright infringers. Against google's alleged monopoly of online information, only a very large sum of money to invest in a

competing search engine can offset the market that google now controls in search traffic and search advertisements. There is no "fair use" exception to google's alleged monopoly over information that would balance the playing field.

In years past, the anti-trust branch of the Department of Justice may have tried to break up a monopoly and/or have a governmental agency regulate the resulting company(ies). For example, in Maryland, the Public Service Commission is responsible for watchdogging the utility and phone companies. Verizon, which operates in several states including Maryland, is a smaller version of AT&T from the 1970s (or perhaps larger given the overall growth in telecommunications in the U.S. over the last thirty years).

The question to be answered is whether Verizon is any more responsive to customers today than AT&T was before the big break up, and whether Verizon is any less stifling of competition and innovation than its predecessor, AT&T. Answering these questions may help to answer whether google ought, as a matter of policy, to be broken into smaller operating groups and/or regulated by the federal government like an "internet utility" company.

With regards to responsiveness, this is a hard question to answer. A regulated utility like Verizon is still a very large entity, and as a matter of statistics, Verizon will make a substantial number of errors in service provision and billing that will lead to user complaints. I don't have any hard data on complaints over time or resolution rates to compare pre- and post-break up of the entity. And as to google, I'm not sure that this is much of an issue. The truth is that there are other search engines in the market today, and it is very easy for an internet user to access these search engines. They may not have the same content indexed, but all of them use some form of search advertising to help subsidize your ability to freely search on them (or you have to pay a subscription fee to use them). The state of search today may not really compare with the customer service issues of telecom customers of years past that were stuck working with the Baby Bell to get their phone to work properly.

With regards to the problem of stifling competition, the telecommunications bust at the beginning of this century was in part the result of the Baby Bells like Bell Atlantic/Verizon who controlled the last mile infrastructure that connected competing telecoms to customers. After a century, the Baby Bells had so much more invested in the public phone and data networks that no small start up could possibly compete. And whether court-ordered or not, the engineers at Verizon were not going to make a competitor's service request a higher priority than servicing direct Verizon customers.

So, if the equivalent of this is for google to keep its database of indexed books but simply share access to other search engines, I doubt the outcome would be much different - most people trying to find a book would get a better response from google's search engine than a competing search engine. Alternatively, if google were required to publish its search engine algorithms and code, how long and how much money would it take for a competitor to grow to sufficient size to accumulate the scope and depth of data that google now handles every day? Ten years?

Twenty? And would the internet be a better place because there are two identical search engines? This is like when there were two different paper phone books. Other than the

additional tree casualties, I don't think the public was better served with two phone books, and I doubt having two identical databases on the internet of web sites would be much better, either.

What about an Internet Public Service Commission? For the phone company, the PSC in each state is empowered to receive and investigate complaints from customers - typically about a billing problem, but the PSC investigators examine related issues as well. My experiences with the PSC here have been positive. The PSC's opening of an investigation usually gets my complaint to the right person at Verizon, who is then able to resolve the problem that the customer service representative was either not empowered to resolve or unwilling to resolve.

What, then, would the IPSC be charged with handling from the public about google? Lost documents in the internet cloud that were stored with google? Google's search crawler doesn't search my site quickly enough? My web site doesn't show up in search results high enough based on my keywords? Security breaches at google?

I suppose another solution would be to make google a national library and attach it to the Library of Congress, which could use the google revenue stream to pay for scanning and indexing everything in the Library of Congress to make it generally available to the public. Unrelated parts of google could be spun off as private enterprises that would not operate with public money or public regulation (such as the cloud computing aspects of google).

Stay tuned for developments!