Paper by an anonymous student
21W.784 – Aden Evens
Essay 2 Draft
10/12/2005

Digital Texts

Pre-digital information resided mainly in the form of printed texts. Copyright laws regarding the internet as they stand today are shamefully counterintuitive. The fundamental concept of the internet is to be able to communicate and share information quickly and easily. Why then are internet copyright laws, namely the DMCA, designed specifically to prevent such practices? Granted, the nature of digital information and the internet allows for easy unethical reproduction, but it is important to distinguish that there are ethical uses as well, and they are too beneficial to deny.

The ethical uses are well within reach. Google Print, for example, is a project which aims to digitize millions of volumes from libraries across the world. Their entire texts would be instantaneously searchable, creating a sort of universal knowledge base. However, as legislation exists today, only works published before 1923 fall under public domain. This is absurd, given that anyone could walk into a public library and read anything he wants. All Google is doing is essentially reducing the time it takes to get the information in front of your eyes. Deliberately depriving the public of such a large knowledge base when it already exists is ludicrous.

The first lawsuit involving Google Print was filed September 21, 2005. The Authors Guild is trying to file a class-action lawsuit on the grounds that they as copyright holders were not consulted before the inclusion of their works in the project.

Texts from books archived by Google are delivered in the form of images. This means that text cannot be easily copied and pasted into the computer's "clipboard" for easy reproduction. It provides essentially the same experience as going to a library or book store and casually flipping through the book. In fact, the experience is lessened in that through Google's service, the consumer can only see a few pages surrounding the text that he searched for. It is not possible to go directly to a desired page, or a page at random. But as law stands today, the authors are very much entitled to consent before inclusion. This is giant obstacle in the development of this technology, as well as in the ethics of the law.

Yahoo has announced its plans about a project similar to Google Print, but they will digitize only those works whose copyright holders consent. This is undeniably the most law-respecting method, but the process seems like a huge hassle. Google claims their methods fall under "fair use," but fair use is very loosely defined, as it is a matter of ethics. There are some very general rules and various court decisions, but qualifications are decided case by case. This is very inefficient. Google has put its effort on hiatus until November 2005 to allow time for any authors whose works are scheduled to be digitized to opt out of the process. This seems to be "fair" enough, but the Authors Guild still wants the owner of each work to be contacted. While this is very frustrating to proponents of the service, it is hard to argue under current law that this is unreasonable. The problem is that the internet is being subjected to pre-digital ethical standards. The internet calls for an ethical paradigm shift, but this has not been accepted.

The internet raises many issues about ethics that have never been seen before. One reason for this is the nature of the internet and digital media. There are two major

pre-digital forms of widely distributed information: radio and television. Both are very similar in that the content providers have control. The record industry is paid royalties and the television networks can pause their information for advertisements. With the internet however, the consumer is and always will be in control of what information he receives. The internet however is not a one-way broadcast. It is highly non-standardized and interactive. Consumers are going to use the internet no matter what, so content providers have no choice but to cater to this new technology. A simple digitization of works and a payment method for the entire work though will not suffice. Texts must be presented in an innovative and user-friendly manner, and Google is already there.

The name "Google Print" may be misleading to some authors, though. "Print" might make it seem like Google is trying to publish works online, which is not the case. There is a reason the e-book phenomenon has not caught on in the mainstream. People still want to hold a tangible book when reading. Consequently, making the process of finding a book easier benefits not only the consumer, but also the author. Many large publishing companies have already submitted books to Google for scanning and archiving in response to Google's invitation (Wyatt). Google Print is a search and reference tool, not some corporate loophole to read books for free.

This is true for fiction (non-reference) books, but what about nonfiction reference books? Surely any researcher would first consult a tool like Google Print. It has the benefits of an internet search – virtually unlimited information instantly – and the certainty of reliable sources traditionally found only in brick-and-mortar libraries. And frankly, this *should* be the first step for research. Currently journals and periodicals are searchable through existing services. Sites like Wikopedia offer full-fledged free online

encyclopedias. Sure, authors deserve compensation for their work, but could this be the early signs of the reference book becoming obsolete? Maybe the next evolutionary step is an online pay service. If so, consumers are not going to go to many different publisher's sites; a central database would be necessary. Ideally all non-threatening, non-private/personal information in the world should be free, but this would leave no incentive for the creation of new knowledge. Advertisements go a long way – Google's non-obtrusive text based ads could easily compensate non-fiction authors and leave the end-user free access to information (Wyatt). A non-fiction author's intention is to spread knowledge and get recognized for his contributions, and nothing makes this easier if the information is free.

Parallels may be drawn to going to the book store and looking at a book, but copyright advocates would argue that the reader is that much closer to actually purchasing the book. Google however leaves 3 or 4 links to buy the book online, which is possible with 2 clicks of the mouse. It is very clear that Google's intentions are not to make money off of copyrighted works – that would be unethical. Google will profit because it's service is superior, not because they are displaying information freely available at libraries and bookstores.

Google's basic internet search engine displays snippets of web sites, which are all copyrighted. They also place ads on every search page, making money in the process. However the websites gain as well with increased traffic. Theoretically Google's Print venture could prove to be similarly beneficial to authors. It would undoubtedly lead to more people stumbling upon a given work, and also increased sales.

To display anything on a computer from the internet, the source bits, whatever they may be (text, audio, video) must be reproduced. This alone is evidence enough that copyright laws need to be drastically redefined for the internet. Internet users expect instant access to information – blame it on the technology, but this is reality – so old ethical standards have no place in cyberspace. There are ways for consumers to have free information and authors to be compensated, but this requires a new and separate set of legislation. The internet is too valuable to waste.

Robert S Boynton. "Righting Copyright: Fair Use and 'Digital Environmentalism.'"

   Bookforum. (February/March 2005).

Edward Wyatt. "Writers Sue Google, Accusing It of Copyright Violation" [online:web],

updated 21 September 2005, cited 12 October 2005. URL:

   http://www.nytimes.com/2005/09/21/technology/21book.html?ex=1284955200&

   en=0fac2a8661f7deba&ei=5090&partner=rssuserland&emc=rss