Spring 3-11-2005

Don't Get Goggle-Eyed Over Google's Plan to Digitize

Mark Y. Herring
Winthrop University, herringm@winthrop.edu

The Chronicle Review

March 11, 2005

Don't Get Goggle-Eyed Over Google's Plan to Digitize

By MARK Y. HERRING

If you haven't heard about Google's plan to digitize millions of books, you must live in another galaxy. Hardly a news outlet in the country, digitized or no, missed the story at the end of last year.

Most people were pleased by the news. It seemed that books would finally be available at your fingertips. Google had embarked on a grand scheme to digitize the world's greatest works, in cooperation with the world's greatest libraries. Break out the champagne!

Not a few bean counters at colleges and universities around the world must have thought, "At long last. We can kick the library in the archives and be done with that financial black hole." Some librarians may have had a similar vision of the future and been dismayed, although most of them were optimistic about Google's plan.

Digitization is big news; it's a good idea; and it's inevitable. But let's not get all goggle-eyed over Google right away. Here are five reasons not to tear up your library card quite yet.

Copyright. A recent Chronicle article ("Google Will Digitize and Search Millions of Books From 5 Top Research Libraries," January 7) was one of the few early reports to mention copyright. Current copyright law, to say nothing of Congress's continuing interest in increasing the length of time that works are protected by copyright, should give everyone at Google heartburn.

At least in the early days, Google plans to rely chiefly on books that are in the public domain -- in general, works published before 1923 -- to avoid paying substantial fees to the copyright holders. The company says that for more-recent books it will provide only a few short excerpts, which it claims would not violate copyright. However, some publishers argue that scanning a book to digitize it constitutes reproduction, for which permission is required by law.

That permission can be expensive. Only recently the library where I work encountered the sting of copyright fees on a small scale, when we asked permission to digitize an article from a book. The publisher charged us the same price as it would have if we had been putting together a whole course pack. The rationale was probably that digitization is replacing course packs, or that previous copyright fees were too low.

Of course the giant Google will have far more influence over publishers than any one library could. But will Google have so much influence as to make copyright fees too low to matter? That's doubtful.

Past failures. Four other companies have tried to do just the sort of digitization that Google is undertaking, and they have had problems. One of them, NetLibrary, filed for Chapter 11 bankruptcy and became, in much-scaled-down form, a subsidiary of the Online Computer Library Center. A second, Questia, remains independent, but it has reduced its work force significantly and is on shaky legs. Each company began with over $80-million in venture capital; neither found that to be enough. The third company, Project Xanadu, is little more than a Web presence now.

Project Gutenberg has done better than the other companies, but it still begs for dollars on its Web site from would-be users, foundations, and anyone else who might lend an ear, and a dime. As a free site for digitized texts (http://www.promo.net/pg/index.html), it specializes in public-domain books and includes some of the works that Google is interested in. But Google also wants to digitize books still in copyright. Does it have a better idea than the previous for-profit businesses, or at least more money to make an old idea work this time around?

Preserving books. Yes, the new machines that Google has can digitize pages with incredible speed. But no matter how fast (and faster, in this case, may not be better), digitization is not good for books, however good it may be for the reader. Who is going to pay for the books damaged in digitization? What happens when a rare book is damaged?

My guess is that Google has underestimated, perhaps substantially, the percentage of books that will be damaged or that cannot undergo rapid digitization. Not only will some books be too fragile, or bound too tightly to lie flat, but even some newer books, owing to rapid manufacture, fall out of their bindings in 12 months or less. Handling -- even by careful digitizers -- will doubtless leave more than a few volumes without covers. Working with both groups of titles will increase Google's costs.

Google's future. What would happen to all the digitized books if -- perish the thought -- Google's scheme comes to an unhappy end, like NetLibrary, Questia, or Xanadu?

It would be very easy for libraries to become overreliant on Google, with pressures on them from every side to reduce costs. In that case, what would librarians do if Google suddenly vanished or went out of the digitization business?

Ecological concerns. Whenever any of us arrives at a Web site that has information we need, what do we usually do after checking out the first or second screen? We hit the print button.

Imagine thousands of students, faculty and staff members, and other library patrons all punching that print key. Of course Google wouldn't pay for the printing. But even if the libraries that offer access to Google's digitized material pass the printing costs on to their patrons, will our glorious digitized library come at the expense of the few forests we have left?

Other concerns also come to mind. For example, what about the increased potential for plagiarism? What about Google's heavy reliance on material in English? The head

of the National Library of France has expressed his worry that the project will be "powerfully marked by the view of Anglo-Saxons" (The Chronicle, March 4).

What kind of advertising will Google use to pay at least some of the costs of digitization? Academics tend to be particularly allergic to ads and other distractions on their computer screens. Google already relies on ads to cover its costs; presumably it will do the same for digitization. Would scholars tolerate having an ad about, say, erectile dysfunction pop up as they read Stanley Fish's Surprised by Sin, in order to have the work digitized?

The digitized "library" would undeniably be for picking and choosing, not really for reading. Is that the attitude toward books that we want to encourage -- the view that sound bites are more important than substantive thought?

Those are not necessarily insurmountable obstacles for Google. However, they are formidable.

Besides, the portability, convenience, and even comfort of a book are integral components of our intellectual lives. No one has yet made a convincing case that it's time to give up on books -- or libraries.

Mark Y. Herring is dean of library services at Winthrop University. His most recent book is Raising Funds With Friends Groups (Neal-Schuman, 2004).