



Online books

Pulp friction

CAMBRIDGE, MASSACHUSETTS, LONDON AND NEW YORK

Internet companies are racing to get books online, but publishers are understandably wary

INFORMATION wants to be free," according to a celebrated aphorism from the early days of the internet. Yet this ethos has been creating new headaches recently. As search-engine firms and others unveil plans to place books online, publishers fear that the services may end up devouring their business, either: by bypassing them or because the initiatives threaten to make their copyrights redundant.

This month, Google announced that it is moving forward with its plans to digitise books from several big libraries, despite two lawsuits filed in October by authors and publishers who claim that the firm's actions violate their copyrights. (Google says its actions are legal under a "fair use" exception in the law.) Meanwhile, Yahoo! and Microsoft have unveiled plans to make books accessible online through the Open Content Alliance, and Microsoft has struck a deal with the British Library to digitise 100,000 books in its collection that are no longer under copyright. Not to be outdone, Amazon said it plans to sell online access to books on a "pay-per-view" basis. And Random House, the world's largest publisher of trade books, last week became the first big publishing house to move into online book access.

Creating a virtual Library of Alexandria has long been a dream of techies and

book-lovers alike. Project Gutenberg, a digitisation initiative dating back to the 1970s, currently boasts over 17,000 books in around 45 languages. This summer, European nations backed a "digital library" plan to place literary works online. For readers, the idea of being able to access the aggregate knowledge of humanity on a single device from anywhere in the world seems a benefit of mythic proportions, and the obvious next step in man's quest—from stone slabs to papyrus to movable type to ethereal digital bits—to document the world in words and symbols.

"In the future, the only thing that will get read is something that will be online. If it isn't online, it doesn't exist", proclaims Jim Gerber, director of content partnerships at Google. Yet for search companies such as Google, Yahoo! and Microsoft's MSN, digitising books is particularly important because it represents the next critical phase of the industry. Most of the web has already been scanned and indexed—there are already between 8 billion and 10 billion items online. Although search technology is constantly tweaked to provide better performance and more relevant results, studies by Microsoft have shown that around half of all search queries fail to provide the information that users want. "We need to get offline content online. Off-

line is where trusted content is, and where people who need to answer questions go," explains Danielle Tiedt, manager of search content acquisition at MSN. "Books are only the first step," she says.

The cost of digitising a book can be as low as 10 cents per page, and as much as \$100 per book if it is longer, oversized, or requires more effort (though for recent titles, digital files already exist). Business models to recoup the investment remain embryonic for now, but include paid advertisements at the bottom of a web page, charging a few cents to view a page and possibly subscription pricing.

Unbundling books

But digitising and indexing books does more than shift printed material online: it permits the "unbundling" of content from the printed page and creates new access and sales opportunities. For example, you might not have to buy a celebrity chef's entire book of recipes before a dinner party—but instead just download the one you want. And with electronic distribution, print-on-demand technology and the notion of "the long tail"—the idea that niche items, not just blockbusters, can be lucrative—books may stop going "out of print".

However, the nub is how the money flows. Nigel Newton, the boss of Bloomsbury (the publishing house that discovered J.K. Rowling, author of the "Harry Potter" books), says that if publishers let technology firms digitise their content this will eventually dilute their control over it, especially as literary content becomes increasingly sold and consumed in digital form. "The publisher's job is to publish, and that includes electronically," he says.

Publishers are determined to make sure ►

that they will benefit from the digitisation of their product. They could certainly do with any extra growth it may bring: consumer-book publishing is one of the most mature industries in media. The total market for all books worldwide totalled \$107.5 billion last year. But the American market grew by only 1.8% a year between 1999 and 2004, according to Veronis Suhler Stevenson, a bank. Moreover, the amount of time that people spend on books is falling gradually, to the benefit of television, the internet and video games.

So far, the internet has been a double-edged sword for book publishers. On one hand, online sales are growing at 8-9%, much faster than at book shops. But at the same time, the internet is facilitating massive growth in sales of used books, from which authors and publishers get nothing. Second-hand book sales represented 1% of trade sales in 2002 in America, according to PricewaterhouseCoopers, but by 2004 had grown to almost a fifth.

Publishers admit that the entry of Google and other tech firms has galvanised them to pay attention to digitisation. "The fact is that Google's and Amazon's actions have stimulated the energy for this to take off," says Ian Hudson, group managing director of Random House in London. "Otherwise we would have dragged on for ages working it out." Now they are focused closely on the issue of when digitising books requires the permission of the copyright holders, along with a payment.

Most of the initiatives avoid the issue by digitising only works that can be freely reproduced. MSN is initially scanning only books from the British Library that are out of copyright. In other cases, firms are striking deals with publishers for the right to

provide online access. This is how Amazon can offer its long-standing "search inside this book" feature, as well as its plan to sell access to pages for a fee. As for Random House, its content will be sold on a pay-per-page basis through third parties like search engines and booksellers (though pages cannot be saved, printed or copied).

Google, however, is playing it both ways. While it does seek permission from publishers for certain books, it believes that digitising libraries does not require permission from the copyright holder, under a "fair use" provision that allows excerpts to be used freely. Google considers the service simply an enhanced card catalogue system.

Trade associations representing publishers and authors are suing Google, claiming that the very act of scanning books without permission is an illegal reproduction. The case promises to keep the lawyers busy. Google seems to have begun back-pedalling, noting that the books it is currently scanning are ones that are out of copyright. It is even working on a model of pay-per-view charging, according to one publishing executive.

Few dispute that the services will be a boon for the public and that books will eventually go digital, though the technology for portable display devices is still in its infancy. "As an industry we should embrace the opportunity that that will bring," says Mr Newton. In time, once robust economic models can be worked out, they will benefit authors and publishers too. In 1987 Stewart Brand, a technology pundit, said, in the very next breath after his oft-quoted aphorism: "Information also wants to be expensive." But no one seems to remember that bit. ■