

December 18, 2004

Questions and Praise for Google Web Library

By FELICIA R. LEE

When Randall C. Jimerson, the president of the Society of American Archivists, heard of Google's plan to convert certain holdings at Oxford University and at some of the leading research libraries in the United States into digital files, freely searchable over the Web, he asked, "What are they thinking?"

Mr. Jimerson had worries. Who would select the material? How would it be organized and identified to avoid mountains of decontextualized excerpts? Would Google users eventually forgo the experience of holding a book, actually seeing a historical document, the serendipity of slow research?

But in recent interviews, many scholars and librarians applauded the announcement by Google, the operator of the world's most popular Internet search service, to digitize some of the collections at Oxford, the University of Michigan, Stanford and the New York Public Library.

The plan, in the words of Paul Duguid, information specialist at the University of California at Berkeley, will "blast wide" open the walls around the libraries of world-class institutions.

David Nasaw, an historian and director of the Center for the Humanities at the City University of New York's Graduate Center, said the ability to use keywords to locate books and documents could save academics travel time and money and ease and broaden the scope of their research.

But Google's plan - which many saw as the first step toward creating a global virtual library - has huge implications for information gathering and use, also raising concerns among those interviewed.

No one forecast a brave new world without actual libraries. Rather, they raised questions.

How will research be improved for students already struggling with, among other things, how to authenticate Internet information? What new roles will librarians play in helping people parse a vast amount of more easily obtainable information? Will libraries have to cooperate to prevent redundancy in their collections?

Each agreement with a library is slightly different. Google plans to digitize nearly all the eight million books in Stanford's collection and the seven million at Michigan. The Harvard project will initially be limited to only about 40,000 volumes. The scanning at Bodleian Library at Oxford will be limited to an unspecified number of books published before 1900, while the New York Public Library project will involve fragile material not under copyright that library officials said would be of interest primarily to scholars.

"This all captures people's imagination in a wonderful way," said Kate Wittenberg, director of the Electronic Publishing Initiative at Columbia University. "But whether it's right or wrong is not the whole question and not the whole answer."

This year Ms. Wittenberg's group completed a three-year study of research habits that included 1,233 students across the country. The study concluded that electronic resources had become the main tool for gathering information, particularly among undergraduates.

But, "What I've learned is that libraries help people formulate questions as well as find answers," Ms. Wittenberg said. "Who will do that in a virtual world?"

On the other hand, she said, an enhanced databank could make it easier for students to research topics across disciplines, changing the questions that professors ask and providing more robust answers. For example, a topic like "climate change" touches on both political science and science, she said, and "in the physical world, the books about them are in two different buildings at Columbia."

And because many students have trouble recognizing reliable Web sources, it cannot hurt that reputable libraries have only published, peer-reviewed materials, Ms. Wittenberg said. But looking ahead, she wondered about the vast amounts of materials and original documents housed outside libraries, in museums and archives.

Mr. Jimerson said, "A scanned image will only tell you some things, and the sheer volume of records makes scanning everything difficult." But he added that he supported Google's plan in theory. "I recall the story of a gentleman being in a library and watching a researcher sniff books," he said. "It turned out that the aroma of vinegar was still embedded in those that had been treated with vinegar to prevent cholera during an epidemic."

Likewise, Robert Darnton, a professor of history at Princeton who is writing a book about the history of books, noted that by looking at a book's binding and paper quality, a researcher can discern much about the period in which it was published, the publisher and the intended audience.

"There may be some false consciousnesses about this breakthrough, that all learning will be at our fingertips," Mr. Darnton said of the plans to enhance Google's database. He saw

room for both Google and real-world research.

Some interviewed were concerned that Google could not fully reproduce material that was still under copyright protection, which means all books published in the United States after 1923. And in this day and age, Mr. Nasaw said, far too many students already read excerpts and seldom read the full texts.

It is not an either-or situation. Libraries have already been changed by the Internet, said Paul LeClerc, president and chief executive of the New York Public Library. Libraries will still be needed to collect, classify and store information, he said.

"TV did not replace radio," Mr. LeClerc said, "Videos and and DVD's did not replace people going to the movies. It's still easier to read a book by hand than online."

"The New York Public Library Web site gets three-fourths of a billion hits a year from 200 different countries and territories, and that's with no marketing or advertising," he said. "That's the context in which this new element has to be placed."

"We had 13 million reader visits last year," he continued. "We're serving a multiplicity of audiences - we serve people physically and virtually. It's an enormous contribution to human intellectual development."

Carol Brey-Casiano, president of the American Library Association, forecast a time when local libraries would more closely reflect the needs of their regions and population rather than all offering the same books. It means that libraries will become more collaborative, she said.

Already, libraries buy fewer reference materials because such materials are online, she said. At the same time, the number of library visitors doubled in the last 10 years to 1.2 billion visits a year now, she added, with many visitors seeking help in managing vast amounts of information. As she put it: "People are saying, 'I went on Google and I got 40,000 hits. Now what?' "

Many university leaders realize that for most people, information does not exist unless it is online, said Paul Courant, provost and executive vice president for academic affairs at the University of Michigan.

He said many universities wanted to digitize their holdings and wondered about collaborating on buying books to avoid redundancy in an increasingly digital world. Google's plan answered their needs, he said.

Mr. Courant acknowledged that some people in Michigan's library system were worried about losing the distinctiveness of the university's collection, as part of a vast database. But he theorized that Google would work as a form of advertisement to lead more people

to the libraries' doors.

"The librarians will also continue to be responsible for archiving and curating our own material and collecting it," he said. "Many scholars will go online and say, 'I have to go see the book' and come here." Mr. Courant envisioned that in 20 years huge archives would be shared by institutions. While the world needs "tens of thousands of copies of 'To the Lighthouse,' " he said, "we don't need to have a zillion copies of some arcane monograph written by a sociologist in 1951."

[Copyright 2004 The New York Times Company](#) | [Home](#) | [Privacy Policy](#) | [Search](#) | [Corrections](#) | [RSS](#) | [Help](#) | [Back to Top](#)