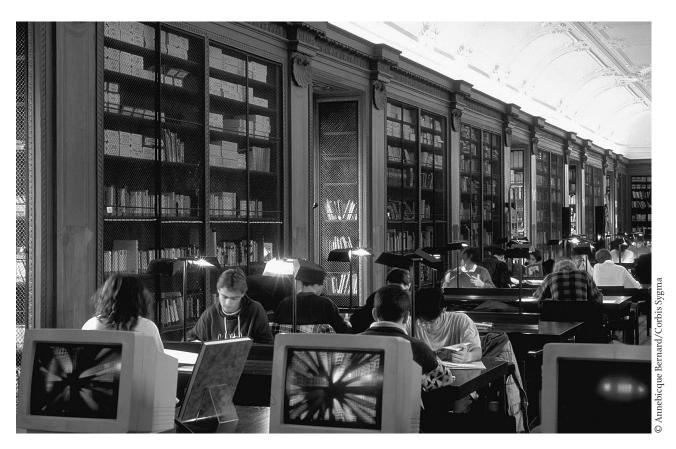# Academy Meetings

The Research Library in the Digital Age

Robert Darnton

This presentation, the fourth S. T. Lee Lecture in the Humanities, was given at the 1925th Stated Meeting, held at the House of the Academy on March 13, 2008. Bernard Bailyn, Adams University Professor Emeritus at Harvard University and a Fellow of the Academy, introduced Robert Darnton.

Robert Darnton

Robert Darnton is Carl H. Pforzheimer University Professor at Harvard University and Director of the Harvard University Library. He has been a Fellow of the American Academy of Arts and Sciences since 1980.

Information is exploding so furiously around us and information technology is changing at such bewildering speed that we face a fundamental problem: How do we orient ourselves in the new landscape? What, for example, will become of research libraries in the face of technological marvels such as Google? How do we make sense of it all?

I have no answer to that problem, but I can suggest an approach to it: look at the history of the ways information has been communicated. Simplifying things radically, you could say that there have been four fundamental changes in information technology since humans learned to speak.

Somewhere around 4000 BC humans learned to write. Egyptian hieroglyphs go back to about 3200 BC, alphabetical writing to 1000 BC. According to scholars like Jack Goody, the invention of writing was the most important technological breakthrough in the history of humanity. It transformed mankind's relation to the past and opened a way for the emergence of the book as a force in history.

The history of books led to a second technological shift when the codex replaced the scroll sometime soon after the beginning of the Christian era. By the third century AD, the codex – that is, books with pages that

you turn as opposed to scrolls that you roll – became crucial to the spread of Christianity. It transformed the experience of reading: the page emerged as a unit of perception, and readers were able to leaf through a clearly articulated text, one that eventually included differentiated words (that is, words separated by spaces), paragraphs, and chapters, along with tables of contents, indexes, and other reader's aids.

The codex, in turn, was transformed by the invention of printing with movable type in the 1450s. To be sure, the Chinese developed movable type around 1045, and the Koreans used metal characters rather than wooden blocks around 1230. But Gutenberg's invention, unlike those of the Far East, spread like wildfire, bringing the book within the reach of ever-widening circles of readers. The technology of printing did not change for nearly four centuries, but the reading public grew larger and larger, thanks to improvements in

What will become of research libraries in the face of technological marvels such as Google?

literacy, education, and access to the printed word. Pamphlets and newspapers, printed by steam-driven presses on paper made from wood pulp rather than rags, extended the process of democratization so that a mass public came into existence during the second half of the nineteenth century.

The fourth great change, electronic communication, took place yesterday, or the day before, depending on how you measure it. The Internet dates from 1974, at least as a term. It developed from ARPANET, which went back to 1969, and from earlier experiments in communication among networks of computers. The Web began as a means of communication among physicists in 1991. Websites and search engines became common in the mid-1990s. And from that point everyone knows the succession of brand names that have made electronic communication an everyday experience: Gopher, Mosaic, Netscape, Internet Explorer, and Google, which was founded in 1998.

When strung out in this manner, the pace of change seems breathtaking: from writing to the codex, 4,300 years; from the codex to movable type, 1,150 years; from movable type to the Internet, 524 years; from the Internet to search engines, 17 years; from search engines to Google's algorithmic relevance ranking, 7 years; and who knows what is just around the corner or coming out the pipeline?

Each change in the technology has transformed the information landscape, and the speedup has continued at such a rate as to seem both unstoppable and incomprehensible. In the long view – what French historians call *la longue durée* – the general picture looks quite clear – or, rather, dizzying. But by aligning the facts in this manner, I have made them lead to an excessively dramatic conclusion. Historians, American as well as French, often play such tricks. By rearranging the evidence, it is possible to arrive at a different picture, one that emphasizes continuity instead of change. The continuity I have in mind has to do with the nature of information itself, or, to put it differently, the inherent instability of texts. In place of the long-term view of technological transformations, which underlies the common notion that we have just entered a new era, the information age, I want to argue that every age was an age of information, each in its own way, and that information has always been unstable.

Let's begin with the Internet and work backward in time. More than a million blogs have emerged during the last few years. They have given rise to a rich lore of anecdotes about the spread of misinformation, some of which sound like urban myths. But I believe the following story is true, though I can't vouch for its accuracy, having picked it up from the Internet myself. As a spoof, a satirical newspaper, *The Onion*, printed a story about an architect who had created a new kind of building in Washington, D.C., one with a convertible dome. On sunny days, you could push a button, the dome would roll back, and the building would look like a football stadium. On rainy days it looks like Congress. The story traveled from website to website until it arrived in China, where it was printed in the *Beijing Evening News*. Then it was taken up by *The Los Angeles Times, The San Francisco*

There have been four fundamental changes in information technology since humans learned to speak: the invention of writing, the codex, movable type, and the Internet.

Chronicle, Reuters, CNN, Wired.com, and countless blogs as a story about the Chinese view of the United States: they think we live in convertible buildings, just as we drive around in convertible cars.

Other stories about blogging point to the same conclusion: blogs create news, and news can take the form of a textual reality that trumps the reality under our noses. Today many reporters spend more time tracking blogs than they do checking out traditional sources, such as the spokespersons of public authorities. News in the information age has broken loose from its conventional moorings, creating possibilities of misinformation on a global scale. We live in a time of unprecedented accessibility to information that is increasingly unreliable. Or do we?

I would argue that news has always been an artifact and that it never corresponded exactly to what actually happened. We take today's front page as a mirror of yesterday's events, but it was made up yesterday evening – literally, by "make-up" editors, who designed page one according to arbitrary conventions: lead story on the far-right column, off-lead on the left, soft news inside or below the fold, features set off by special kinds of headlines. Typographical design orients the reader and shapes the meaning of the news. News itself takes the form of narratives composed by professionals according to conventions that they picked up in the course of their training – the "inverted pyramid" mode of exposition, the "color" lead, the code for "high" and "the highest" sources, and so on. News is not what happened but a story about what happened.

Of course, many reporters do their best to be accurate, but they must conform to the

conventions of their craft, and there is always slippage between their choice of words and the nature of an event as experienced or perceived by others. Ask anyone involved in a reported happening. They will tell you that they did not recognize themselves or the event in the story that appeared in the paper. Sophisticated readers in the Soviet Union learned to distrust everything that appeared in *Pravda* and even to take nonappearances as a sign of something going on. On August 31, 1980, when Lech Walesa signed the agreement with the Polish government that created Solidarity as an independent trade union, the Polish people refused at first to believe it, not because the news failed to reach them but because it was reported on the state-controlled television.

We live in a time of unprecedented accessibility to information that is increasingly unreliable.

I used to be a newspaper reporter myself. I got my basic training as a college kid covering police headquarters in Newark in 1959. Although I had worked on school newspapers, I did not know what news was – that is, what events would make a story and what combination of words would make it into print after passing muster with the night city editor. When events reached headquarters, they normally took the form of "squeal sheets" or typed reports of calls received at the central switchboard. Squeal sheets concerned everything from stray dogs to murders, and they accumulated at a rate of a dozen every half hour. My job was to collect them from a lieutenant on the second floor, go through them for anything that might be news, and announce the potential news to the veteran reporters from a dozen papers playing poker in the press room on the ground floor. The poker game acted as a filter for the news. One of the reporters would say if something I selected would be worth checking out. I did the checking, usually by phone calls to key offices like the homicide squad. If the information was good enough, I would tell the poker game, whose members would phone it in to their city desks. But it had to be really good – that is, what ordinary people would

consider bad – to warrant interrupting the never-ending game. Poker was everyone's main interest – everyone but me: I could not afford to play (cards cost a dollar ante, a lot of money in those days), and I needed to develop a nose for news.

I soon learned to disregard DOAs (dead on arrival, meaning ordinary deaths) and robberies of gas stations, but it took time for me to spot something really "good," like a hold-up in a respectable store or a water main break at a central location. One day I found a squeal sheet that was so good – it combined rape and murder – that I went straight to the homicide squad instead of reporting first to the poker game. When I showed it to the lieutenant on duty, he looked at me in disgust: "Don't you see this, kid?" he said, pointing to a *B* in parentheses after the names of the victim and the suspect. Only then did I notice that every name was followed by a *B* or a *W*. I did not know that crimes involving black people did not qualify as news.

Having learned to write news, I now distrust newspapers as a source of information, and I am often surprised by historians who take them as primary sources for knowing what really happened. I think newspapers should be read for information about how contemporaries construed events, rather than for reliable knowledge of the events themselves. A study of news during the American Revolution by a graduate student of mine, Will Slauter, provides an example. Will followed accounts of Washington's defeat at the Battle of Brandywine as it was refracted in the American and European press. In the eighteenth century, news normally took the form of isolated paragraphs rather than "stories" as we know them now, and newspapers lifted most of their paragraphs from each other, adding new material picked up from gossips in coffeehouses or ship captains returning from voyages. A loyalist New York newspaper printed the first news of Brandywine with a letter from Washington informing Congress that he had been forced to retreat before the British forces under General William Howe. A copy of the paper traveled by ship, passing from New York to Halifax, Glasgow, and Edinburgh, where the paragraph and the letter were reprinted in a local newspaper.

The Edinburgh reprints were then reprinted in several London papers, each time under-

New information technology should force us to rethink the notion of information itself . . . . Instead of firmly fixed documents, we must deal with multiple, mutable texts.

going subtle changes. The changes were important because speculators were betting huge sums on the course of the American war, while bears were battling bulls on the Stock Exchange, and the government was about to present a budget to Parliament, where the pro-American opposition was threatening to overthrow the ministry of Lord North. At a distance of 3,000 miles and four to six weeks of travel by ship, events in America were crucial for the resolution of this financial and political crisis.

What had actually happened? Londoners had learned to mistrust their newspapers, which frequently distorted the news as they lifted paragraphs from each other. That the original paragraph came from a loyalist American paper made it suspect to the reading public. Its roundabout route made it look even more doubtful, for why would Washington announce his own defeat, while Howe had not yet claimed victory in a dispatch sent directly from Philadelphia, near the scene of the action? Moreover, some reports noted that Lafayette had been wounded in the battle, an impossibility to British readers, who believed (wrongly from earlier, inaccurate reports) that Lafayette was far away from Brandywine, fighting against General John Burgoyne near Canada.

Finally, close readings of Washington's letter revealed stylistic touches that could not have come from the pen of a general. One – the use of "arraying" instead of "arranging" troops – later turned out to be a typographical error. Many Londoners therefore concluded the report was a fraud, designed to promote the interests of the bull speculators and the Tory politicians – all the more so as the press coverage became increasingly inflated through the process of plagiarism. Some London papers claimed that the minor defeat had been a major catastrophe for the

Americans, one that had ended with the annihilation of the rebel army and the death of Washington himself. (In fact, he was reported dead four times during the coverage of the war, and the London press declared Benedict Arnold dead 26 times.)

Le Courrier de l'Europe, a French newspaper produced in London, printed a translated digest of the English reports with a note warning that they probably were false. This version of the event passed through a dozen French papers produced in the Low Countries, the Rhineland, Switzerland, and France itself. By the time it arrived in Versailles, the news of Washington's defeat had been completely discounted. The Comte de Vergennes, France's foreign minister, therefore continued to favor military intervention on the side of the Americans. And in London, when Howe's report of his victory finally arrived after a long delay (he had unaccountably neglected to write for two weeks), it was eclipsed by the more spectacular news of Burgoyne's defeat at Saratoga. So the defeat at Brandywine turned into a case of miswritten and misread news – a media *non*-event whose meaning was determined by the process of its transmission, like the blogging about the convertible dome and the filtering of crime reports in Newark's police headquarters.

Information has never been stable. That may be a truism, but it bears pondering. It could serve as a corrective to the belief that the speedup in technological change has catapulted us into a new age, in which information has spun completely out of control. I would argue that the new information technology should force us to rethink the notion of information itself. It should not be understood as if it took the form of hard facts or nuggets of reality ready to be quarried out of newspapers, archives, and libraries, but rather as messages that are constantly being reshaped in the process of transmission. Instead of firmly fixed documents, we must deal with multiple, mutable texts. By studying them skeptically on our computer screens, we can learn how to read our daily newspaper more effectively – and even how to appreciate old books.

Bibliographers came around to this view long before the Internet. Sir Walter Greg developed it at the end of the nineteenth century, and

Donald McKenzie perfected it at the end of the twentieth century. Their work provides an answer to the questions raised by bloggers, Googlers, and other enthusiasts of the World Wide Web: Why save more than one copy of a book? Why spend large sums to purchase first editions? Aren't rare book collections doomed to obsolescence now that everything will be available on the Internet? Unbelievers used to dismiss Henry Clay Folger's determination to accumulate copies of the First Folio edition of Shakespeare as the mania of a crank. The First Folio, published in 1623, seven years after Shakespeare's death, contained the earliest collection of his plays, but most collectors assumed that one copy would be enough for any research library. When Folger's collection grew beyond three dozen copies, his friends scoffed at him as Forty Folio Folger. Since then, however, bibliographers have mined that collection for crucial information, not only for editing the plays but also for performing them.

> *To students in the 1950s, libraries looked like citadels of learning. Knowledge came packaged between hard covers, and a great library seemed to contain all of it.*

They have demonstrated that 18 of the 36 plays in the First Folio had never before been printed. Four were known earlier only from faulty copies known as "bad" quartos – booklets of individual plays printed during Shakespeare's lifetime, often by unscrupulous publishers using corrupted versions of the texts. Twelve were reprinted in modified form from relatively good quartos; and only two were reprinted without change from earlier quarto editions. Since none of Shakespeare's manuscripts has survived, differences between these texts can be crucial in determining what he wrote. But the First Folio cannot simply be compared with the quartos, because every copy of the Folio is different from every other copy. While being printed in Isaac Jaggard's shop in 1622 and 1623, the book went through three very different issues. Some copies lacked *Troilus and Cres-*

sida, some included a complete *Troilus,* and some had the main text of *Troilus* but without its prologue and with a crossed-out ending to *Romeo and Juliet* on the reverse side of the leaf containing *Troilus*'s first scene.

The differences were compounded by at least 100 stop-press corrections and by the peculiar practices of at least nine compositors who set the copy while also working on other jobs – and occasionally abandoning Shakespeare to an incompetent teenage apprentice. By arguing from the variations in the texts, bibliographers like Charlton Hinman and Peter Blayney have reconstructed the production process and thus arrived at convincing conclusions about the most important works in the English language. This painstaking scholarship could not have been done without Mr. Folger's Folios.

Of course, Shakespeare is a special case. But textual stability never existed in the pre-Internet eras. The most widely diffused edition of Diderot's *Encyclopédie* in eighteenth-century France contained hundreds of pages that did not exist in the original edition. Its editor was a clergyman who padded the text with excerpts from a sermon by his bishop in order to win the bishop's patronage. Voltaire considered the *Encyclopédie* so imperfect that he designed his last great work, *Questions sur l'Encyclopédie,* as a nine-volume sequel to it. In order to spice up his text and to increase its diffusion, he collaborated with pirates behind the back of his own publisher, adding passages to the pirated editions.

In fact, Voltaire toyed with his texts so much that booksellers complained. As soon as they sold one edition of a work, another would appear, featuring additions and corrections by the author. Their customers protested. Some even said that they would not buy an edition of Voltaire's complete works – and there were many, each different from the others – until he died, an event eagerly anticipated by retailers throughout the book trade.

Piracy was so pervasive in early-modern Europe that best sellers could not be blockbusters as they are today. Instead of being produced in huge numbers by one publisher, they were printed simultaneously in many small editions by many publishers, each racing to make the most of a market unconstrained by copyright. Few pirates attempted to produce accurate counterfeits of the orig-

Students today still respect their libraries, but reading rooms are nearly empty on some campuses.

inal editions. They abridged, expanded, and reworked texts as they pleased, without worrying about the authors' intentions. They behaved as deconstructionists *avant la lettre*.

The issue of textual stability leads to the general question about the role of research libraries in the age of the Internet. I cannot pretend to offer easy answers, but I would like to put the question in perspective by discussing two views of the library, which I would describe as grand illusions – grand and partly true.

To students in the 1950s, libraries looked like citadels of learning. Knowledge came packaged between hard covers, and a great library seemed to contain all of it. To climb the steps of the New York Public Library, past the stone lions guarding its entrance and into the monumental reading room on the third floor, was to enter a world that included everything known. The knowledge came ordered into standard categories that could be pursued through a card catalogue and into the pages of the books. In colleges everywhere the library stood at the center of the campus. It was the most important building, a temple set off by classical columns where one read in silence: no noise, no food, no disturbances beyond a furtive glance at a potential date bent over a book in quiet contemplation.

Students today still respect their libraries, but reading rooms are nearly empty on some campuses. In order to entice the students back, some librarians offer them armchairs for lounging and chatting, even drinks and snacks, never mind about the crumbs. Modern or postmodern students do most of their research at computers in their rooms. To them, knowledge comes online, not in libraries. They know that libraries could never contain it all within their walls, because information is endless, extending everywhere on the Internet, and to find it one needs a search engine, not a card catalogue. But this, too, may be a grand illusion – or, to put it positively, there is something to be said for both visions, the library as a citadel and the Internet as

open space. We have come to the problems posed by Google Book Search.

In 2006 Google signed agreements with five great research libraries – the New York Public, Harvard, Michigan, Stanford, and Oxford's Bodleian – to digitize their books. Books in copyright posed a problem, which soon was compounded by lawsuits from publishers and authors. But putting that aside, the Google proposal seemed to offer a way to make all book learning available to all people, or at least those privileged enough to have access to the World Wide Web. It promised to be the ultimate stage in the democratization of knowledge set in motion by the invention of writing, the codex, movable type, and the Internet.

Now, I speak as a Google enthusiast. I believe Google Book Search really will make book learning accessible on a new, worldwide scale, despite the great digital divide that separates the poor from the computerized. It also will open up possibilities for research involving vast quantities of data, which could never be mastered without digitization. As an example of what the future holds, I would cite the Electronic Enlightenment, a project sponsored by the Voltaire Foundation of Oxford. By digitizing the correspondence of Voltaire, Rousseau, Franklin, and Jefferson – about 200 volumes in superb, scholarly editions – it will, in effect, re-create the trans-Atlantic republic of letters from the eighteenth century. The letters of many other philosophers, from Locke and Bayle to Bentham and Bernardin de Saint-Pierre, will be integrated into this database, so that scholars will be able to trace references to individuals, books, and ideas throughout the entire network of correspondence that undergirded the Enlightenment.

Many such projects – notably American Memory sponsored by the Library of Congress[1] and The Valley of the Shadow created at the University of Virginia[2] – have demonstrated the feasibility and usefulness of databases on this scale. But their success does not

prove that Google Book Search, the largest undertaking of them all, will make research libraries obsolete. On the contrary, Google will make them more important than ever. To support this view, I would like to organize my argument around eight points.

First, according to the most utopian claim of the Googlers, Google can put virtually all printed books online. That claim is misleading, and it raises the danger of creating false consciousness, because it may lull us into neglecting our libraries. What percentage of the books in the United States – never mind the rest of the world – will be digitized by Google: 75 percent? 50 percent? 25 percent? Even if the figure is 90 percent, the residual, nondigitized books could be important. I recently discovered an extraordinary libertine novel, *Les Bohémiens,* by an unknown author, the Marquis de Pelleport, who wrote it in the Bastille at the same time that the Marquis de Sade was writing his novels in a nearby cell.

There is something to be said for the library as a citadel and the Internet as open space.

I think that Pelleport's book, published in 1790, is far better than anything Sade produced; and whatever its aesthetic merits, it reveals a great deal about the condition of writers in pre-Revolutionary France. Yet only six copies of it exist, as far as I can tell, none of them available on the Internet.[3] (The Library of Congress, which has a copy, has not opened its holdings to Google.) If Google missed this book, and other books like it, the researcher who relied on Google would never be able to locate works of great importance.

The criteria of importance change from generation to generation, so we cannot know what will matter to our descendants. They may learn a lot from studying our harlequin novels or computer manuals or telephone books. Literary scholars and historians today depend heavily on research in almanacs, chapbooks, and other kinds of "popular" lit-

---

[1] It is, according to the site, "a digital record of American history and creativity," including sound recordings, prints, maps, and many images.

[2] An archive of letters, diaries, official records, periodicals, and images documenting the life of two communities – one Northern, one Southern – 200 miles apart in the Shenandoah Valley during the years 1859 – 1870.

[3] See my essay, "Finding a Lost Prince of Bohemia," *The New York Review of Books,* April 3, 2008, 44 – 48.

erature, yet few of those works from the seventeenth and eighteenth centuries have survived. They were printed on cheap paper, sold in flimsy covers, read to pieces, and ignored by collectors and librarians who did not consider them "literature." A researcher at Trinity College, Dublin, recently discovered a drawer full of forgotten ballad books, each one the only copy in existence, each priceless in the eyes of the modern scholar, though it had seemed worthless two centuries ago.

Second, although Google pursued an intelligent strategy by signing up five great libraries, their combined holdings will not come close to exhausting the stock of books in the United States. Contrary to what one might expect, there is little redundancy in the holdings of the five libraries: 60 percent of the books being digitized by Google exists in only one of

No computer screen gives satisfaction like the printed page. But the Internet delivers data that can be transformed into a classical codex.

them. There are about 543 million volumes in the research libraries of the United States. Google reportedly set its initial goal of digitizing at 15 million. As Google signs up more libraries – at last count, 28 are participating in Google Book Search – the representativeness of its digitized database will improve. But it has not yet ventured into special collections, where the rarest works are to be found. And of course the totality of world literature – all the books in all the languages of the world – lies far beyond Google's capacity to digitize.

Third, although it is to be hoped that the publishers, authors, and Google will settle their dispute, it is difficult to see how copyright will cease to pose a problem. According to the copyright law of 1976 and the copyright extension law of 1998, most books published after 1923 are currently covered by copyright, and copyright now extends to the life of the author plus 70 years. For books in the public domain, Google will allow readers to view

the full text and print every page. For books under copyright, however, Google will display only a few lines at a time, which it claims is legal under fair use. Google may persuade the publishers and authors to surrender their claims to books published between 1923 and the recent past, but will it get them to modify their copyrights in the present and future? In 2006, 291,920 new titles were published in the United States, and the number of new books in print has increased nearly every year for the last decade, despite the spread of electronic publishing. How can Google keep up with current production while at the same time digitizing all the books accumulated over the centuries? Better to increase the acquisitions of our research libraries than to trust Google to preserve future books for the benefit of future generations. Google defines its mission as the communication of information – right now, today; it does not commit itself to conserving texts indefinitely.

Fourth, companies decline rapidly in the fast-changing environment of electronic technology. Google may disappear or be eclipsed by an even greater technology, which could make its database as outdated and inaccessible as many of our old floppy disks and CD-ROMs. Electronic enterprises come and go. Research libraries last for centuries. Better to fortify them than to declare them obsolete, because obsolescence is built into the electronic media.

Fifth, Google will make mistakes. Despite its concern for quality and quality control, it will miss books, skip pages, blur images, and fail in many ways to reproduce texts perfectly. Once we believed that microfilm would solve the problem of preserving texts. Now we know better.

Sixth, as in the case of microfilm, there is no guarantee that Google's copies will last. Bits become degraded over time. Documents may get lost in cyberspace, owing to the obsolescence of the medium in which they are encoded. Hardware and software become extinct at a distressing rate. Unless the vexatious problem of digital preservation is solved, all texts "born digital" belong to an endangered species. The obsession with developing new media has inhibited efforts to preserve the old. We have lost 80 percent of all silent films and 50 percent of all films made

Don't think of the library as a warehouse or a museum. While dispensing books, most research libraries operate as nerve centers for transmitting electronic impulses.

before World War II. Nothing preserves texts better than ink imbedded on paper, especially paper manufactured before the nineteenth century, except texts written in parchment or engraved in stone. The best preservation system ever invented was the old-fashioned, premodern book.

Seventh, Google plans to digitize many versions of each book, taking whatever it gets as the copies appear, assembly-line fashion, from the shelves; but will it make all of them available? If so, which one will it put at the top of its search list? Ordinary readers could get lost while searching among thousands of different editions of Shakespeare's plays, so they will depend on the editions that Google makes most easily accessible. Will Google determine its relevance ranking of books in the same way that it ranks references to everything else, from toothpaste to movie stars? It now has a secret algorithm to rank Web pages according to the frequency of use among the pages linked to them, and presumably it will come up with some such algorithm in order to rank the demand for books. But nothing suggests that it will take account of the standards prescribed by bibliographers, such as the first edition to appear in print or the edition that corresponds most closely to the expressed intention of the author. Google employs hundreds, perhaps thousands, of engineers, but, as far as I know, not a single bibliographer. Its innocence of any visible concern for bibliography is particularly regrettable in that most texts, as I have just argued, were unstable throughout most of the history of printing. No single copy of an eighteenth-century best seller will do justice to the endless variety of editions. Serious scholars will have to study and compare many editions, in the original versions, not in the digitized reproductions

that Google will sort out according to criteria that probably will have nothing to do with bibliographical scholarship.

Eighth, even if the digitized image on the computer screen is accurate, it will fail to capture crucial aspects of a book. For example, size. The experience of reading a small duodecimo, designed to be held easily in one hand, differs considerably from that of reading a heavy folio propped up on a book stand. It is important to get the feel of a book – the texture of its paper, the quality of its printing, the nature of its binding. Its physical aspects provide clues about its existence as an element in a social and economic system; and if it contains margin notes, it can reveal a great deal about its place in the intellectual life of its readers.

Books also give off special smells. According to a recent survey of French students, 43 percent consider smell to be one of the most important qualities of printed books – so important that they resist buying odorless electronic books. CaféScribe, a French online publisher, is trying to counteract that reaction by giving its customers a sticker that will give off a fusty, bookish smell when it is attached to their computers.

When I read an old book, I hold its pages up to the light and often find among the fibers of the paper little circles made by drops from the hand of the vatman as he made the sheet – or bits of shirts and petticoats that failed to be ground up adequately during the preparation of the pulp. I once found a fingerprint of a pressman enclosed in the binding of an eighteenth-century *Encyclopédie* – testimony to tricks in the trade of printers, who sometimes spread too much ink on the type in order to make it easier to get an impression by pulling the bar of the press.

I realize, however, that considerations of "feel" and "smell" may seem to undercut my argument. Most readers care about the text, not the physical medium in which it is embedded; and by indulging my fascination with print and paper, I may expose myself to accusations of romanticizing or of reacting like an old-fashioned, ultra-bookish scholar who wants nothing more than to retreat into a rare book room. I plead guilty. I love rare book rooms, even the kind that make you put on gloves before handling their treasures.

Rare book rooms are a vital part of research libraries, the part that is most inaccessible to Google. But libraries also provide places for ordinary readers to immerse themselves in books, quiet places in comfortable settings, where the codex can be appreciated in all its individuality.

In fact, the strongest argument for the old-fashioned book is its effectiveness for ordinary readers. Thanks to Google, scholars are able to search, navigate, harvest, mine, deep link, and crawl (the terms vary along with the technology) through millions of websites

> *As a citadel of learning and as a platform for adventure on the Internet, the research library still deserves to stand at the center of the campus, preserving the past and accumulating energy for the future.*

and electronic texts. At the same time, anyone in search of a good read can pick up a printed volume and thumb through it at ease, enjoying the magic of words as ink on paper. No computer screen gives satisfaction like the printed page. But the Internet delivers data that can be transformed into a classical codex. It already has made print-on-demand a thriving industry, and it promises to make books available from computers that will operate like ATM machines: log in, order electronically, and out comes a printed and bound volume. Perhaps some day a text on a handheld screen will please the eye as thoroughly as a page of a codex produced two thousand years ago.

Meanwhile, I say: shore up the library. Stock it with printed matter. Reinforce its reading rooms. But don't think of it as a warehouse or a museum. While dispensing books, most research libraries operate as nerve centers for transmitting electronic impulses. They acquire data sets, maintain digital reposito-

ries, provide access to e-journals, and orchestrate information systems that reach deep into laboratories as well as studies. Many of them are sharing their intellectual wealth with the rest of the world by permitting Google to digitize their printed collections. Therefore, I also say: long live Google, but don't count on it living long enough to replace that venerable building with the Corinthian columns. As a citadel of learning and as a platform for adventure on the Internet, the research library still deserves to stand at the center of the campus, preserving the past and accumulating energy for the future. ∎