

bookselling environment, then book sales are generally promoted. If book excerpts are displayed in a search engine's advertising-driven environment, then ad sales are generally promoted.

33. To the extent Google's unauthorized displays of books encourages readers to search at its ad-supported search engine, rather than logging in to Amazon's retail environment, Google is hurting the sales of authors' books. For this reason, and many, many others, authors and other rights holders should have control of when their books are copied in their entirety, and where their books are displayed.

34. Google, in other words, disrupted the commercial, permission-driven development of book-search-and-display at online bookstores in order to gain a competitive advantage over other search engines. In the process, it distributed millions of digitized books to universities, placing those books beyond the control of authors and publishers and putting them at plain risk of widespread infringement.

I declare under penalty of perjury that the foregoing is true and correct.



Paul Aiken

Dated: New York, New York
August 26, 2013

Press Releases

[<< Back](#)

Amazon.Com Launches "Search Inside the Book" Enabling Customers to Discover Books by Searching and Previewing the Text Inside

Leading-Edge Search Technology Delivers a Book-Buying Experience Previously Unavailable in the Physical or Online Worlds

SEATTLE--Oct. 23, 2003-- Amazon.com (Nasdaq:AMZN) today announced the launch of its latest innovation for customers, Search Inside the Book, an entirely new way for customers to find and discover books by searching the full text inside them, not just matches to author or title keywords.

In collaboration with publishers, Amazon.com is enabling customers to find books at Amazon.com based on every word inside more than 120,000 books -- more than 33 million pages of searchable text. Customers can also preview the inside text of these books. Search Inside the Book is integrated into Amazon.com's standard search and includes books from all genres.

"Innovation drives customer experience, and Search Inside the Book is a great example," said Jeff Bezos, founder and CEO, Amazon.com. "With the help of publishers, we're offering a completely new way for people to find the books they want."

"The customer in me loves this," said Maureen Egen, president of Time Warner Book Group.

Through Search Inside the Book, customers can find, discover and buy titles from more than 190 publishers, including the industry's largest: Wiley, Time Warner Book Group, Simon & Schuster, Inc., Random House, Inc., Publishers Group West, Incorporated, McGraw-Hill Professional, Holtzbrinck Publishers and HarperCollins Publishers.

Here's how it works:

- Customers interested in resistojet propulsion can search for "resistojet" and will automatically see a list of the many books at Amazon.com that contain this term in the text, in addition to those books that contain the term in the title.
- Books with "resistojet" in the text will display an excerpt including that word or phrase and a link that says "See more references to 'resistojet' in this book." By clicking on this link, customers will see a list of excerpts from all pages on which "resistojet" appears in the book they selected, as well as a link to view the full page from the book on which the excerpt appears.
- Once customers have clicked on the link to a specific page and signed in with their Amazon.com user name and password, they can preview relevant pages, including the page they selected, and search for other terms of interest within the book.

Additional examples include:

- Those looking to explore the scenic Multnomah Falls can search for this term to discover a wide range of related titles ranging from lodging and hiking guidebooks, to Leslie Carraway's "Land Mammals of Oregon."
- Searching for "Curse of the Bambino" provides baseball lovers and Red Sox fans with a variety of books on the subject -- everything from "Deep Change: Discovering the Leader Within" to "Uncle John's Giant 10th Anniversary Bathroom Reader."

- Customers who search for "product price elasticity" will discover more than 100 titles, including Michael Porter's "Competitive Strategy: Techniques for Analyzing Industries and Competitors," and David Besanko's "Economics of Strategy, 2nd Edition," which includes 32 references to this exact subject.

Customers are encouraged to type words and phrases of interest into the new search box, and enjoy whatever surprises they might find.

To encourage customers to share their experiences using Search Inside the Book, Amazon.com is running a contest that will award one grand prize winner a Segway Human Transporter. Ten additional customers will each win a \$100 Amazon.com gift certificate. Visit our "How It Works" page (www.amazon.com/exec/obidos/tg/browse/-/10197021/) to enter the contest, learn more about the new feature, and see more sample searches.

About Amazon.com

Amazon.com, a Fortune 500 company based in Seattle, opened its virtual doors on the World Wide Web in July 1995 and today offers Earth's Biggest Selection. Amazon.com seeks to be Earth's most customer-centric company, where customers can find and discover anything they might want to buy online, and endeavors to offer its customers the lowest possible prices. Amazon.com and sellers list millions of unique new and used items in categories such as apparel and accessories, sporting goods, electronics, computers, kitchenware and housewares, books, music, DVDs, videos, cameras and photo items, toys, baby items and baby registry, software, computer and video games, cell phones and service, tools and hardware, travel services, magazine subscriptions and outdoor living items.

This announcement contains forward-looking statements within the meaning of Section 27A of the Securities Act of 1933 and Section 21E of the Securities Exchange Act of 1934. Actual results may differ significantly from management's expectations. These forward-looking statements involve risks and uncertainties that include, among others, risks related to potential future losses, significant amount of indebtedness, competition, commercial agreements and strategic alliances, seasonality, potential fluctuations in operating results and rate of growth, foreign exchange rates, management of potential growth, system interruption, international expansion, consumer trends, inventory, fulfillment center optimization, limited operating history, government regulation and taxation, fraud, and new business areas. More information about factors that potentially could affect Amazon.com's financial results is included in Amazon.com's filings with the Securities and Exchange Commission, including its Annual Report on Form 10-K for the year ended December 31, 2002, and all subsequent filings.

CONTACT:

Amazon.com, Seattle
Public Relations, 206-266-7180

SOURCE: Amazon.com



The Great Library of Amazonia

120,000 fully searchable texts and counting ... Jeff Bezos is building the world's biggest digital book archive. It's an info-age dream come true - and the best way to sell books ever.

By Gary Wolf

The fondest dream of the information age is to create an archive of all knowledge. You might call it the Alexandrian fantasy, after the great library founded by Ptolemy I in 286 BC. Through centuries of aggressive acquisition, the librarians of Alexandria, Egypt, collected hundreds of thousands of texts. None survives. During a final wave of destruction, in AD 641, invaders fed the bound volumes and papyrus scrolls into the furnaces of the public baths, where they are said to have burned for six months. "The lesson," says Brewster Kahle, founder of the Internet Archive, "is to keep more than one copy."

Kahle recently gave a copy of his digital archive of 10 billion Web pages to a new library in Alexandria. On a visit to the city last year, he sat down with Suzanne Mubarak, the wife of Egypt's president, and discussed his gift, which has all the advantages of a modern electronic resource: It can be instantly updated, easily searched, and endlessly replicated. Mubarak, with diplomatic politeness, allowed that she was impressed. Still, she ventured a protest: "But I love books!"

Therein lies a problem. Books are an ancient and proven medium. Their physical form inspires passion. But their very physicality makes books inaccessible to the multi-terabyte databases of modern Alexandrian projects. Books take time to transport. Their text vanishes and their pages yellow in a rash of foxing. Most important, it's still shockingly difficult to find information buried in books. Even as the Internet has revived hope of a universal library and Google seems to promise an answer to every query, books have remained a dark region in the universe of information. We want books to be as accessible and searchable as the Web. On the other hand, we still want them to be books.

An ingenious attempt to illuminate the dark region of books is under way at Amazon.com. Over the past spring and summer, the company created an unrivaled digital archive of more than 120,000 books. The goal is to quickly add most of Amazon's multimillion-title catalog. The entire collection, which went live Oct. 23, is searchable, and every page is viewable.

To build the archive, Amazon CEO Jeff Bezos has had to unravel a tangle of technological and copyright problems. His solution promises to remake the publishing business and give Amazon a powerful new weapon in its battle against online competitors such as Yahoo, Google, and eBay. But the most interesting thing about the archive is the way it resolves the paradox of the book, respecting its physical form while transcending its limits.

I recently drove to a home in Silicon Valley and spent a few hours digitally searching the text of books. My host was Udi Manber, an Israeli-born computer scientist and author of a popular textbook, *Introduction to Algorithms: A Creative Approach*. Ten years ago, while developing a seminal piece of Unix search software called *agrep*, Manber came up with a concept for an

information tool he has yet to build. It was supposed to search the mess of papers on his desk. The idea that you could perform a digital search of physical objects has long fascinated him. "Why not have users take pictures of their bookshelf?" Manber asked when we first met. "We could scan the images, extract the titles, and then let them search the entire text of the books they own."

The notion of Amazon scanning all of its books but allowing users to search only those they own is a clever way around the central barrier to creating a digital archive: Copyrights are distributed among tens of thousands of publishers and authors. But when Manber told Bezos his idea, he found the Amazon founder ready to work on a grander scale. Bezos wanted his customers to be able to search everything.

In his small, ranch-style Palo Alto house, Manber and I sit side by side at a table near the kitchen as he begins typing my queries into his laptop. The computer is connected to a prototype of the archive, which at the time of my visit is scheduled to go live in a few weeks. Within seconds, I am captivated. The experience reminds me of how I felt a decade ago, when I first began browsing the Web. Back then, the Web was still small, and most of my time was spent peeking into the homepages of physicists and engineers. Even so, the power of the new network was unmistakable. The thrill didn't come from the content of the pages but from the structure of the Web itself, its obvious scalability and ease of navigation.

Amazon's new archive is more densely populated than the early Web was, but it's still far from complete. With its 120,000 titles, the archive has about as many books as a big brick-and-mortar store. Still, this is plenty to create a familiar sensation of vertigo as an expansive new territory suddenly opens up.

The more specific the search, the more rewarding the experience. For instance, I've recently become interested in Boss Tweed, New York's most famous pillager of public money. Manber types "Boss Tweed" into his search engine. Out pop a few books with *Boss Tweed* in the title. But the more intriguing results come from deep within books I never would have thought to check: *A Confederacy of Dunces*, by John Kennedy Toole; *American Psycho*, by Bret Easton Ellis; *Forever: A Novel*, by Pete Hamill. I immediately recognize the power of the archive to make connections hitherto unseen. As the number of searchable books increases, it will become possible to trace the appearance of people and events in published literature and to follow the most digressive pathways of our collective intellectual life.

From the Hamill reference, I link to a page in the afterward on which he cites books that influenced his portrait of Tweed. There, on the screen, is the cream of the research performed by a great metropolitan writer and editor. Some of the books Hamill recommends are out of print, but all are available either new or used on Amazon.

With persistence, serendipity and plenty of time in a library, I may have found these titles myself. The Amazon archive is dizzying not because it unearths books that would necessarily have languished in obscurity, but because it renders their contents instantly visible in response to a search. It allows quick query revisions, backtracking, and exploration. It provides a new form of map.

Getting to this point represents a significant technological feat. Most of the material in the archive comes from scanned pages of actual books. This may be surprising, given that most books today are written on PCs, e-mailed to publishers, typeset on computers, and printed on digital presses. But many publishers still do not have push-button access to the digital files of the books they put out. Insofar as the files exist, they are often scattered around the desktops of editors, designers, and contract printers. For books more than a few years old, complete digital files may be lost. John Wiley & Sons contributed 5,000 titles to the Amazon project -- all of them in physical form.

Fortunately, mass scanning has grown increasingly feasible (see "[3 Ways to Scan a Library](#)"), with the cost dropping to as low as \$1 each. Amazon sent some of the books to scanning centers in low-wage countries like India and the Philippines; others were run in the United States using specialty machines to ensure accurate color and to handle oversize volumes. Some books can be chopped out of their bindings and fed into scanners, others have to be babied by a human, who turns pages one by one. Remarkably, Amazon was already doing so much data processing in its regular business that the huge task of reading the images of the books and converting them into a plain-text database was handled by idle computers at one of the company's backup centers.

The copyrights to these titles are spread among countless owners. How was it possible to create a publicly accessible database from material whose ownership is so tangled? Amazon's solution is audacious: The company simply denies it has built an electronic library at all. "This is not an ebook project!" Manber says. And in a sense he is right. The archive is intentionally crippled. A search brings back not text, but pictures -- pictures of pages. You can find the page that responds to your query, read it on your screen, and browse a few pages backward and forward. But you cannot download, copy, or read the book from beginning to end. There is no way to link directly to any page of a book. If you want to read an extensive excerpt, you must turn to the physical volume -- which, of course, you can conveniently purchase from Amazon. Users will be asked to give their credit card number before looking at pages in the archive, and they won't be able to view more than a few thousand pages per month, or more than 20 percent of any single book.

Manber has built a powerful, even mind-boggling tool, then added powerful constraints. "The point is to help users find a book," says Manber, "not to make a new source of information."

Bezos is vehement on this point. He has sold publishers on the idea that digitizing hundreds of thousands of copyright books won't undermine the conventional bookselling business. "It is critical that this be understood as a way to get publishers and authors in contact with customers," he says in an interview at Amazon's Seattle headquarters. "We're perfectly aligned with these folks. Our goal is to sell more books!"

Bezos has some good evidence to back up his argument. Amazon has consistently added features that have proven to increase book sales. Through its customer reviews, used-book business, and personalized recommendations, the company constantly puts its customers in contact with new titles. Amazon is a machine that stimulates the acquisitive urge of readers. It appeals to their specialized interests.

It makes people buy books. But Amazon's scheme would never work if users really wanted their books in digital form. The magic of the archive lies in the assumption that physical books are irreplaceable. The electronic text is simply an enhancement of the physical object.

The Amazon project -- dubbed Search Inside the Book -- represents a bold step toward the dream of a universal library. Bezos refuses any such allusion. But outlines of the Alexandrian fantasy can clearly be made out in Amazon's innocent book-purchasing tool. The company's success at launching a massive archive of digital books will undoubtedly fuel enthusiasm for overturning the current publishing and copyright regime.

I first talked with Brewster Kahle a dozen years ago over a plate of execrable spaghetti in the kitchen of a flat in San Francisco's Mission District. The apartment served as the headquarters of WAIS, one of the first Internet search engines, and Kahle was sharing a dinner with me and several of his employees. He was in his twenties then -- thin, with a mass of unruly curly hair, a rapid manner of speech, and an unguarded expression. Kahle was already one of the great enthusiasts of universal information access. A few months earlier, he had left his job at Thinking Machines, the legendary builder of massively parallel supercomputers, and devoted himself full-time to refining and selling WAIS.

To Kahle, it was obvious that vast amounts of useful material could be shared with the general public via computers, but the Web did not yet exist, and most of the major databases were not linked. You couldn't do a comprehensive search. WAIS was meant as a remedy, and it proved a modest success. However, its most significant contribution to Kahle's evangelical mission was a byproduct of the stock market bubble: In the spring of 1995, AOL bought WAIS for \$15 million in stock. AOL stock soared, and Kahle became rich.

With his money, Kahle started the Internet Archive, while also creating another company that offered a clever Web search tool called Alexa. In 1999, as the bubble continued to expand, Alexa was sold to Amazon for \$250 million in stock, and Kahle became richer. He's now committed to public service. The computers of the Internet Archive are in a Mission district warehouse. The headquarters are in a ramshackle house in the Presidio, a decommissioned Army base near the Golden Gate Bridge. The office is one of those classic engineer-idealist domains, where programmers go up the fire stairs because the inside stairs are broken, and age-old pizza molders in the refrigerator.

When I call Kahle to ask if I can come talk to him about the state of digital libraries, he says, "Sure, I'm free right now." I find him substantially the same. "What's the average lifespan of a Web page?" Kahle asks me when we meet, and then answers himself: "One hundred days!" He has a slightly accusatory tone, as if I share in the general neglect that led to the erasure of history - or would, if the machines owned by the Internet Archive weren't so busily preserving it.

The goal of the archive is to save digital information and make it accessible to all. But what, exactly, is digital information? As the entertainment industry has learned to its great chagrin, digital information might take the form of music, or movies, or even books. To Kahle, this is a good thing. The products of human knowledge ought not to be squirreled away where they can't ever be found.

Kahle hates the idea that when people think of information, they think only of what's accessible via Google. "Seventy-one percent of college students use the Internet as their research tool of first resort," he says, citing figures from a 2001 PEW Internet Study. "Personally, I think this number is low. For most students today, if something is not on the Net, it doesn't exist."

And yet most books are not on the Net. This means that students, among others, are blind to the most important artifacts of human knowledge. For many students, the Internet actually contracts the universe of knowledge, because it makes the most casual and ephemeral sources the most accessible, while ignoring the published books. "It's shameful," Kahle continues, "because we have the tools to make all books available to everybody. You need three things. Technically, you need storage and connectivity. Storage is easy. For under \$10 million, you can store all published works of humankind back to the Sumerian tablets. The last time they tried this was in Alexandria, and they had an innovative storage mechanism, too. They had papyrus, and papyrus was astonishing compared to clay tablets. But we can do better than the Alexandrians, because we also have connectivity. I have traveled in Uganda and in rural Kenya and seldom been more than one day's walk from an Internet café. It is technologically possible for most kids in the world to have access to all the books in the world."

The third item on Kahle's list has nothing to do with technological know-how; it's simply political will. Here, he finds the situation mixed. "We live in an open society in which the concept of widespread knowledge is embraced as a goal of governance," he says. "Just look at our libraries. Public libraries spend \$7.6 billion a year; academic libraries spend another \$5 billion." That's the good news. The budgets are hard evidence of a public commitment to the Alexandrian ideal. But on the other hand, almost none of this money goes to digitizing books.

The Internet Archive turned Kahle into an expert in managing huge databases of publicly accessible information. Now - in partnership with Carnegie Mellon University, the National Science Foundation, and the governments of India and China - his goal is to create a digital

archive of 1 million books. Books from the US are packed into containers and shipped to India to be scanned and proofread, then the digital files go to the Internet Archive and the books are returned to the owners. Kahle and his partners are hoping to have about 100,000 online by the end of the year, making this project almost as big, at least numerically, as Amazon's effort. "We chose a million books because it's a big number," admits Kahle. "It's something you can strive for."

But in reality, the Million Book Project lags far behind Amazon's effort. For one thing, libraries have been slow to lend parts of their collections. And even then, the project concentrates on digitizing those that are out of copyright. Libraries and nonprofits don't have much leverage with publishers, and since the goal is fully readable online text, there is no system to protect the interests of copyright holders. As a result, many of the titles being digitized by the Million Book Project are government documents, old texts, and books from India and China, where copyright laws are less stringent.

Kahle is happy to sidestep the problem of digitizing commercially successful books. He has no wish to antagonize the publishing industry. What he hates is that the Million Book Project cannot legally digitize countless books that aren't generating money for anybody. US libraries hold about 30 million unique volumes. No one knows how many of those books continue to be protected by copyright or are available from commercial publishers. Still, Kahle says, "they can't be digitized because the copyrights can't be cleared, and the copyrights can't be cleared because it's too much work to identify the copyright holders. Some people call them abandonware. I call them orphans."

"Amazon is taking a cut at the commercially available titles," continues Kahle. "We are going for the public domain titles. But who is taking care of the orphans? Nobody."

This is no longer true. Kahle's plea on behalf of orphaned books, stripped of its sentiment and restated in the rational voice of finance, exactly expresses the logic of Amazon's Alexandrian venture. Latent within the new archive is a business model for selling books that, with a little legal help, ought to vanquish orphanhood forever.

The publishing industry has made great strides since the Roman era. Movable type was invented in 11th-century China, then reinvented in 1450 in Germany. In 1886, Ottmar Mergenthaler created an automatic typesetting machine. In 1983, we got desktop publishing. But publishers continue to edit books using four colors of pencil, and the idea of freely accessible digital files conjures nightmares of a peer-to-peer disaster among media corporations. Things are even going backward - Barnes & Noble recently announced it would stop selling ebooks.

In this context of change, confusion, and fear, Jeff Bezos is forced to behave like a politician. Talk of a universal library elicits no enthusiasm from him. When I mention it, he counsels caution and patience. "You have to start somewhere," he says. "You climb to the top of the first tiny hill, and from there you see the next hill. It's difficult to see what's beyond before you have climbed the first hill."

When I met Bezos the first time, in 1996, there was no masking the radical nature of his ideas. At the time, he was trying hard to prove that you could create a major retail company on the Web. Skeptics abounded, and he answered them with vivid descriptions of the future. All of Amazon's important innovations - starting from the concept of a Web bookstore - have suggested a profound change in the bookselling business, a change that makes it possible to earn a profit by selling a much wider variety of books than any previous retailer, including many titles from the so-called long tail of the popularity curve. "If I have 100,000 books that sell one copy every other year," says Steve Kessel, an Amazon VP, "then in 10 years I've sold more of these, together, than I have of the latest Harry Potter."

In fact, Amazon doesn't have to wade far into the shallows to begin remaking the book business. Books are abandoned by publishers long before their sales are reduced to one copy every other year. Under the current publishing system, a title becomes inefficient at thousands of sales per year. An electronic archive through which readers can find books is an essential counterpart to Bezos' original vision of an infinitely big bookstore, just as Internet search engines are essential to the fragmented, increasingly diverse cultures of the Web.

This vision implies that readers will someday be able to purchase books that are printed at the time they are ordered. On a small scale, that phase of the revolution has already been quietly accomplished. As part of the Million Book Project, Kahle has created an Internet Bookmobile that produces decent-quality paperbacks of out-of-copyright books for about \$1 each. The bookmobile consists of a Ford Windstar minivan with a satellite dish, a computer, a printer, and a binder. Meanwhile, last spring, Amazon announced a partnership with Ingram Industries' Lightning Source subsidiary, a print-on-demand company that offers more than 100,000 titles - with a list that grows by hundreds each week. Lightning Source is the high end, Kahle's Internet Bookmobile is the low end; both operate on the premise that tiny runs of books can be affordably made and sold.

With these tools, the concept of out of print is becoming obsolete. A copyright-friendly archive that allows all books to be easily found plus a books-on-demand printing network gives publishers an economic motive for reactivating entire back catalogs. As for books whose copyright holders cannot be found - Kahle's orphans - this is where the copyright law needs to change. A sensible solution advanced by copyright scholar (and Wired columnist) Lawrence Lessig and written into a bill before Congress requires that copyright be renewed every 50 years for a token sum. Anybody who can't be bothered to pay a dollar or two to hold on to a copyright loses the work to the public domain.

And who will digitize the books once they've been claimed by their copyright holders or lapsed into the public domain? Project Gutenberg, the first book-digitizing initiative, has put about 10,000 titles online and is rapidly accelerating its effort. The Million Book Project, launched by Carnegie Mellon computer scientist Raj Reddy, is eager for more volumes. Lessig, in partnership with Stanford University librarian Michael Keller, will soon announce a free program to digitize any out-of-print book whose copyright holder wants to make it available to the public. And of course, Manber invites copyright holders to offer nonexclusive searching and browsing rights to Amazon, which will digitize the titles and offer access to audiences forever. "Give the books to me," Manber says. "I am glad to do it."

The original vision of a digital archive of all knowledge renounced paper volumes; physical books were seen as antiquated, like papyrus or clay tablets. But if electronic archives prove to raise the value of physical books, a new dream may replace the old one. After talking with Manber, I raise this question with Kevin Kelly, a Wired founding editor who spent part of his summer trying to establish a private cooperative library of digital books. The digital titles in Kelly's library would match the physical books on his shelf. "The idea of ebooks was to do away with paper," he says. "But really, you want to add dimensionality to a physical object rather than take it away. You want an enhanced physical world."

In this enhanced physical world, the logic of the book business is transformed. Human attention is limited, and a massive number of newly browsable books from the long tail necessarily compete with the biggest best-sellers, just as cable siphons audience from the major networks, and just as the Web pulls viewers from TV.

This shifts power away from the people who own finite sets of copyrighted material and toward the people who offer access to information about where this material can be found. Information about books, not ownership of copyrights, becomes a new center of power. Manber is correct when he says that Amazon's Search Inside the Book is not an ebook project. It is merely a

catalog. But a decade of Internet history proves that the catalog is exactly what you want to own.

Of course, Amazon is not merely part of the book business. The Internet as a whole is going through a similar transition. Revenue for Internet companies increasingly comes from users seeking to buy something - from transactions made in the physical world. Even Google, which doesn't sell anything directly, earns most of its cash from advertisers whose messages pop up when users search for info about specific products, such as computers or cars - or books. With retail at the center of the Internet industry, Google is a key competitor because customers begin their online shopping trips at search engines that offer neat algorithms for comparing prices across multiple vendors. Everybody - Yahoo!, eBay, AOL, Microsoft, and, of course, Amazon - wants to be the site of first resort.

All the leading retail sites have better knowledge of their customers than Google. But Google is the leading Internet information tool, period. Google is a window onto the entire Web. On the other hand, the contents of books may be the only publicly accessible data set with the potential to match Google's Web index both for size and utility. Search Inside the Book makes Amazon the sole guide to tens and ultimately hundreds of millions of pages of information. And while Google's business is vulnerable to any competitor that builds a better search engine, Amazon's book archive is the product of negotiated contracts with hundreds of publishers. Amazon has cornered the market on information that was once hidden away in books. The burden of the physical - the fact that the database Amazon uses is linked into a complex system involving real things - gives it a stunning, if perhaps temporary, advantage.

This fall, Amazon announced that it was forming A9.com, a new company devoted exclusively to search technologies. Manber, who leads it, came up with the name by running a simple compression algorithm on the word algorithms. Algorithms begins with A and is followed by nine other letters. When Manber explains the name to me, he notes mischievously that another word can be identically compressed: Alexandria.

Amazon's Alexandrian scheme hinges on the insight that physical books can be turned into electronic databases and then - in the retail process - turned back into physical books. This is one of the boldest maneuvers yet in an intense commercial competition, but for all its cunning, this is a civilized, even civilizing war, one that builds libraries rather than burns them.

3 Ways to Scan a Library

by Dustin Goot

Books today are written on laptops, typeset on PCs, and pumped out on digital presses. But ironically, the most efficient way to create an electronic library is to scan the printed page. The technology has come a long way since the days of the Kurzweil machines - hulking, 1-ton scanner/optical character recognition combos that emerged in the early 1980s. Today, the biggest archiving projects use some combination of these three methods.

Tear Off the Spines

Using a paper guillotine (which looks just like its Bastille cousin), a book's pages are simply lopped off of the binding and sent through a scanner with an automatic page feeder. High-end machines cost \$25,000 and churn through 90 black-and-white pages per minute, front and back. Rare books need not apply.

Ship It Overseas

Workers in India, China, and the Philippines earn about 40 cents an hour to manually turn pages that are zapped by \$15,000 overhead scanners. Carnegie Mellon's Million Book Project alone employs more than 100 Indians for this activity. The Indian government views it as a boon to local employment.

Hire a Robot

Flipping pages is more complicated than meets the eye, but robots are getting the hang of it. Earlier this year, Kirtas Technologies introduced a bot that has both an overhead scanner and an automated page-turning arm. While a book sits open in a special cradle, the arm swoops down, grabs the top page with gentle suction, and turns it. The machine boasts a speed of 1,200 pages an hour, but justifying its six-figure price to frugal librarians is tough.

The Digital Book Brigade

by Erik Malinowski

Amazon's Search Inside the Book project is not the only effort to bring the bookshelf to the desktop - it's just the biggest. From legions of volunteers typing in single pages to small-town contributors translating their personal libraries into myriad languages, a handful of pioneers have blazed the digital book trail.

Initiative	Launched	Books digitized	Goal	What's in it	Who's behind it
Amazon.com's Search Inside the Book	2003	120,000	"Millions"	A cross-section of popular titles	Amazon.com
Project Gutenberg	1971	10,000	1 million	Classics - as long as they're not copyrighted	The nonprofit Literary Archive Foundation
Million Book Project	2001	10,000	1 million	Government docs, old titles, books from India	National Science Foundation, Carnegie Mellon
International Children's Digital Library	2002	262	10,000	Domestic and international children's books	NSF, Institute for Museum and Library Services, University of Maryland
Children's Books Online: The Rosetta Project	2000	119	4,500	Antique children's books from around the world	Nonprofit foundation relying on individual private donations

Contributing editor Gary Wolf (gary@aether.com) is the author of *Wired: A Romance* (Random House).

Copyright © 1993-2004 The Condé Nast Publications Inc. All rights reserved.

Copyright © 1994-2003 Wired Digital, Inc. All rights reserved.

HOME ARCHIVES ABOUT JOHN FOUR LETTER WORDS REBELMOUSE [signup](#) [Enter your email](#) SUBSCRIBE

Thoughts on the intersection of search, media, technology, and more.

SEARCH

Searchblog on     [All Feeds](#)



« [F*cking Spam](#)

[More on A9](#) »

NEWS: A9, AMAZON'S SEARCH PORTAL, GOES LIVE: REVERBERATIONS FELT IN VALLEY

BY [JBAT](#) - APRIL 14, 2004

92

 [0](#)

 [0](#)

0

 [3](#)

0





[A9](#), Amazon's much discussed skunk works search project goes live today, so I can finally write about it. I saw it last month (caveat: unbeknownst to me until recently, Amazon

targeted me as their conduit to break this news – I think they wanted it to move from the blogosphere out, as opposed the WSJ in) and had to keep the damn thing to myself, it was hard, and here's why: On first blush it's a very, very good service, and an intriguing move by Amazon. It raises a clear question: How will Google – and more broadly, the entire search-driven world – react?

My gut tells me the public face will be one of partnership: After all, A9 uses Google's search results and displays at least two paid AdWord listings per result (I've requested comment from Google, you can imagine I'm not the only one...). But I have to wonder: What business is Google in, after all? Is it still in the business of just search – as it was back when it was cutting search provisioning deals right and left, with Yahoo (already ended), AOL (arguable imperiled due to Gmail and other trends), Ask, and Amazon? Is it really still in the business of being an OEM to others, a strategy which allowed it to steal those portals' customers? Or... has it evolved, to a business where it owns a large customer base, one it must now position itself to defend?

It seems to me, Google's position in Amazon's A9 implementation is at best a step backwards. If A9 is as good as it seems to be, every customer that uses and/or switches to A9 becomes an A9 search customer, and, more likely than not, a deeper and far more loyal Amazon customer. (The service incorporates a personal search history and many other really neat tweaks, including a wicked good Toolbar.) In essence, Amazon seems to be making a play for Google's customers. Or it seems that way to me, anyway. Sure, Amazon isn't in the AdWords business. It's happy to outsource that to Google and focus on the entire US retail GDP instead...



[Udi Manber](#), the head of A9 and one of the leading lights of the search community, is understandably evasive when asked about this subject. Google and Amazon have always been friends and partners (despite the fact that "Work at Google" is the top paid link when you search on his name on Google). But as I point out in the introduction to my [Business 2.0 interview](#), to be posted any moment now, one-time partners can quickly become serious competitors in the Search Find Obtain market. And judging from the look of it, A9 is a very direct statement from Amazon: We are now officially in the search business, so get used to it.

One could argue that A9 is a pure commerce play, not a search portal. After all, that's what the folks at Amazon insisted when they founded the company and located it in the heart of Google/YahooLand (ie, Palo Alto). But that argument is disingenuous. First off, take a look at the A9 interface. Where's the commerce? (Answer, it's there, but it's hidden, more on that later when I post on the service itself). And second, I'd argue that you can't really be in the commerce business without having at least a strategy for owning search. The reverse also holds true. It's two ends toward the middle, and by the way, that middle ground is getting damn crowded – AOL, Yahoo, MSN, eBay, IAC, Amazon, Google...

Of course anyone who's been in this game for a while will tell you that the internet industry is rife with cat and mouse games of cooperation turned to competition. Netscape's outsourced its early search traffic to Yahoo, thereby ensuring Yahoo's success. Yahoo paid the favor forward by outsourcing its search to Google, a practice it ended only last quarter. Microsoft built Overture, and crushed Looksmart. And AOL's advertising business is on the rise again, due in large part to a deal with Google, which just announced a [stunning new email service](#) that pretty much decapitates one of AOL's core differentiators (oh, Yahoo and MSN as well...).

What makes this particularly noteworthy is that A9 is built quite literally on top of Google. In short, Amazon has taken the best of Google, and made it, to my mind, a lot better. Sound familiar? Yup, it's what Google did to Yahoo, Yahoo to Netscape...you get the picture.

It all reminds me of a [quote in a recent AP story](#) from Google employee #1:

(The ongoing threat of competition) has helped keep Google from becoming complacent, said Craig Silverstein, the company's director of technology. "If someone should come along and do a better job than us, we know people will switch in a heartbeat."

Something tells me the hearts are beating a bit faster at Yahoo and Google HQs today. Will Google renew its deal with Amazon? Will Bezos and Schmidt put a good face on it and call this a partnership? I have no idea, but man, things are certainly getting interesting in this neck of the woods. More after I talk with folks and get a second order view of the landscape.

(I'll also have a much more complete posting on A9, including a tour of its features and a discussion of its strategic implications later tonight.)

PS- for a tour of what's cool in A9: [Click here.](#)

Related Posts



News: AOL Puts A Stake In The Ground



Layering Commentary Onto Google Earth

Period Ending:	Year Ended:
03/31/04	03/31/03
Revenue	Revenue
1,002	74,823
1,074	7,473
4,290	10,289
10,317	112,746
0	1,270
10,317	112,746
0	1,270
10,317	112,746
0	1,270

Accoona IPO



Reporters Need to Understand Advertising. But Should They Be Making It?



More on A9



Will Google Reintroduce Microblogging?

« [F*cking Spam](#)

[More on A9](#) »

You might also like:



Web 2 - Sergey Stops By



Groupon: That's More Like It



CM Summit: Help Me Interview Dick Costolo



Twitter's Makin' Media



The Web Is Stealing Searches From Microsoft Office

LinkWithin

92 comments

★ 0



Leave a message...

Best Community

Share Settings



Veeresh • 5 years ago

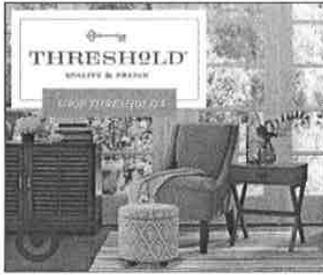
It can be a better solution for an end-user who is in touch with Amazon often in order to search for some product and get the results are mostly relevant to the products.

A9.com lacks some usability and accessibility features. Search results need to be organized well for better access, now it seems little congested content. Some of the features are not compatible with some others browser(Opera). Amazon need to take these bugs severely before end-users start going away.

Thanks!

Veeresh D.

<http://drveresh.googlepages.com>



LIKE SEARCHBLOG

Like

Send

1,172 people like this. Be the first of your friends.

CATEGORIES

Select Category

MOST ACTIVE POSTS

[Is US Culture Veering Toward The Dark and Deadpan?](#)

8 comments · 1 day ago

[Four Minutes on the Future of Marketing](#)

9 comments · 4 days ago

[Great Content, Meet Great Targeting \(And Reach\)](#)

6 comments · 2 weeks ago

TOP POSTS

[Predictions 2012: The Roundup](#)

[Searchblog 2011: The Year In Writing](#)

[Updated: On FareCast: Rip Me Off No More](#)

[The Database of Intentions](#)

FM'S WHITEBOARD

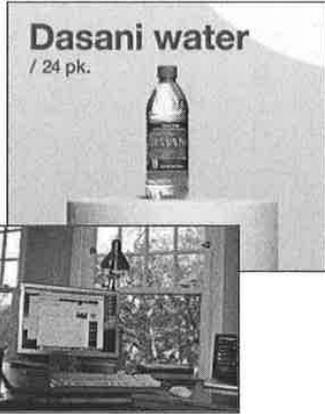
[Federated Media Launching Native and Programmatic Ad Service Blend - Sales and Marketing @ FolioMag.com](#)
August 23, 2013

["We are trying to make it as easy as possible for people who have thoughtful things to say to get..."](#)
August 23, 2013

[Media Agencies Sign on for BuzzFeed's Course in Creating Buzzy Content](#)
August 22, 2013

[Author Spotlight: Anna Beth Chao of AB Chao FM: Where do you...](#)
August 21, 2013

[Colorful, patterned, eccentric socks are trending at FM amongst...](#)
August 20, 2013



A LITTLE ABOUT JOHN

 45, is an entrepreneur, journalist, professor, and author who has founded or co-founded scores of online, conference, magazine, and other media businesses. This work is licensed under a [Creative Commons Attribution - NonCommercial- NoDerivs 2.5 License](#). [Click here to return to the homepage](#) - 2013 John Battelle. | [Terms of Service and Privacy Policies](#)

RELATED LINKS

- [Contact](#)
- [Disclosures](#)
- [Federated Media](#)
- [That Book I Hear He Wrote](#)
- [The Web 2 Conferences](#)

HOME ARCHIVES ABOUT JOHN FOUR LETTER WORDS REBEL MOUSE sign up Enter your email SUBSCRIBE

Thoughts on the intersection of search, media, technology, and more.

SEARCH

Searchblog on     All Feeds



« [Searching Is Getting More Complicated](#) [Good Morning, Your New Net IPO For The Day Is...](#) »

INTERVIEW WITH A9'S MANBER UP ON B2.0

BY JBAT - APRIL 15, 2004

0  0  0  0

BUSINESS2.0 No need to repost it here, as [my interview with Udi](#) is up on Business2.com without subscription walls (thanks 2.0!). Includes an intro which gives an overview of the service and the implications. Good for those of you who don't want to wade through my last two posts on the subject...and gives a bit of insight into how Udi's mind works.

UPDATE: The link apparently is now behind reg, so the column is posted in extended entry. (6/21/04)

]]><![CDATA[

WEB EXCLUSIVE

Can Amazon Unplug Google?

The online retailer is planning to change the way we navigate the Web. An exclusive interview with Udi Manber, whose newly unveiled A9 engine just may drain the juice from the brightest light in search.

By John Battelle, Business2.com, April 15, 2004

Is Amazon (AMZN) taking on Google? Last fall the giant online retailer announced that it was launching a company focused exclusively on search. Lest anyone miss the implications, Amazon promptly rented an office far from its Seattle headquarters — in the epicenter of Silicon Valley, near Stanford University. That's where both Google and Yahoo (YHOO) were born.

Tapped to lead the startup was Udi Manber, Amazon's chief algorithms officer and Yahoo's former chief scientist. Widely respected for his elegant approaches to intractable software problems, Manber was a magnet for Silicon Valley talent. (Either in jest or retaliation, Google responded by buying advertisements on its own search engine for the phrase "Udi Manber"; type his name into the search form and the top ad link offers job opportunities at Google.)

Amazon recently took the wraps off its initial public beta version of A9, and we got the first sneak peek. It's easy to see why the new engine might set Google and other search-driven companies on edge: A9 is a credible step toward making a search engine that knows you and acts as your agent online.

By overlaying what you're looking for onto what you've already found — and, ultimately, what you've consumed — A9 aims for the holy grail of search and e-commerce. While others have tried to create personalized interfaces, none has worked so far. Users don't like to answer questionnaires about what they like and what they don't; often they lie. But Amazon already knows who you are — or at least, what you buy. With A9, the company can factor in what you browse and search for as well.

And the best part, from Amazon's perspective, is that the more things you buy at Amazon, and the more you browse with A9, the better the company will know you. Ideally, it's win-win — you get better results, Amazon gets more sales.

Manber is the first to tell you he's nowhere near his goal of creating the perfect search engine. But A9 — the name stands for the nine letters in the word "algorithm" — is a significant evolution. Blending powerful personalization features with clever interface tweaks and a truly useful toolbar, A9 could well do Google one better. Interestingly, it uses Google's search results — Amazon and Google have a longstanding search deal — but builds on them with an elegant set of features that, taken together, make A9 seem like a new way to search.

In truth, Amazon and Google are speeding from two ends toward a very fertile middle ground where commerce and search meet. When they get there, they'll have formidable company — AOL, eBay (EBAY), Microsoft (MSFT), and Yahoo, to name a few. All of these companies have millions of loyal users and business models built on connecting those who seek with those who sell.

And there's an obvious question: Does Amazon honestly think Google will stand by while Manber creates a better way to search?

It's clear he's just getting started. A native of Israel, Manber has an almost elfish grin and the merry intensity born of someone who clearly knows more than he's allowed to tell. John Battelle asked him a lot of questions anyway.

Walk me through what's different about A9 as a search site.

When people come to A9.com, they have the option to install a browser toolbar and register through Amazon's regular registration process. If you're already an Amazon customer, the site automatically recognizes you. If not, you don't have to register, but if you do, you get the personalized features, and if you install the toolbar, you get even better services.

As to what is new, the most obvious feature is your personal search history — which is integrated into your entire search experience. So your entire search history is available to you, and with the toolbar, that includes all your searches across any search site, as well as all your browsing on the Web. Also, all the interface columns on the site — personal history, or related books, or the Web search column itself — are adjustable and the site remembers the settings the next time you come (the default is that Web search only is opened). I'm very proud of our toolbar's diary feature that allows you to annotate each site you visit. And we have integrated Amazon's "search inside the book" feature into the engine, so now all your results include excerpts from related books.

I understand much of the personalization is made possible by what you call a "history server." What exactly is that?

The history server stores — on our servers — your history of interaction with us for the purpose of bringing that back to you in a very convenient way. Whenever you come to the site, we can show you what you searched for in the past in a very easy-to-organize fashion. If you want to hide some of that, you can opt out at any time. If you install the toolbar, then all your Web browsing, as well as all your searching, is stored as well. And we are working on many different ways to improve that.

So there is a massive repository of data, held on an Amazon server, that tracks where I've been on the Web. Isn't that something of a privacy nightmare?

Our privacy policy is very clear on this subject — we will never share this data with third parties. Also, having the data of what users search for or where they go is not a new concept. Any site that requires registration — AOL or MSN, for example — already has this information. What is new is that we're taking this information and giving it back to the users to make their search experience better and more useful. It gives them more power.

Will A9 ever integrate all the information Amazon already knows about its customers — what they buy, browse, or search — into creating better search results? If I happen to buy a lot of books on jungle cats, will my A9 search results know to serve me the cat-related search results for the query "jaguar" ahead of the car or the football team results?

We have no plans to do this at the moment. It is very hard to do well. If we do this, we want to make sure to do it in a way that makes the search experience work better for users. But again, I can't predict the future. The whole goal here is to get better customer experience. Whatever I can find to get better customer experience, I'll do it.

How can we square the perception that A9 was going to be a commerce engine with the reality of the site as it debuted, which feels much more like the evolution of a pure search site like Google? Will a shopping comparison engine be added soon?

I can't answer that. As an engineer I don't like to talk about things that I haven't yet built. A9 is a way to experiment with several features that help our users. It's a way to bring more users in, and it's a way to see the feedback they give us using those features. We will later use that feedback to improve search both at Amazon and other sites we partner with. It is clear that search is an integral part of any e-commerce site. It's also clear that search is not a solved problem for e-commerce, or beyond e-commerce.

What happens if A9 starts to steal market share from Google?

We don't talk about other companies, and to be honest, we don't even think too much about competitors. We try to concentrate completely on the users. I don't look at it as who we compete with but rather how much something helps the users.

You've said search is not solved. What are the problems that need to be addressed?

In general, the main problem of search is very simple: Too often people do not get the results they need. Solving that is hard. We cannot read people's minds, but we can do much better anticipating their needs and making the process much easier for them.

What might we expect from A9 in the future?

A9 is going to roll out features over time. This is our first beta test — the result of a small team working over a short amount of time. We are growing the group as fast as we can. We have a large queue of very good ideas with the purpose of making search better. Our long-term goal is to improve search. We hope to invent new search technologies and use them on Amazon.com and other partner sites.

John Battelle is a visiting professor at the University of California at Berkeley and author of "The Search" (Portfolio, late 2004).

Find this article at <http://www.business2.com/b2/subscribers/articles/1,17863,611251,00.html>

Related Posts



#5: Nutch Presages a New Kind of Search Engine



Titan Column: Homer



Web 2 Brin Interview Up



Interview: Jonathan Miller, AOL



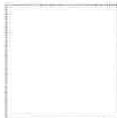
#3: Why Contextual Paid Ads Matter



A Few Questions for Dave Morgan, Founder Tacoda

« [Searching Is Getting More Complicated](#) [Good Morning, Your New Net IPO For The Day Is...](#) »

You might also like:



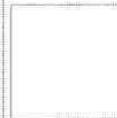
Will Our Industry Ever Innovate Like Morse? Probably Not.



Portrait of Twitter As A Young Media Company



Just In Case Ya Missed It, Google Is Pushing Google+, Hard



Web 2 Summit Points of Control: The Map



Pandora's Facebook Box

LinkWithin

0 comments

★ 0



Leave a message...

Best Community

Share Settings

No one has commented yet.

ALSO ON JOHN BATTELLE'S SEARCHBLOG

What's this?

Great Content, Meet Great Targeting (And Reach)

6 comments • 21 days ago



johnbattelle — Thanks very much!

Four Minutes on the Future of Marketing

9 comments • 21 days ago



johnbattelle — I imagine yes, it will

Is US Culture Veering Toward The Dark and Deadpan?

8 comments • 3 days ago



johnbattelle — Yep, that was what I was commenting on as well..

The Book Lives On

1 comment • 3 months ago



Clint Bagley — Welcome Saral

Comment feed Subscribe via email

Get fast, secure internet access at over 100,000 hotspots.

CLICK HERE TO GET THE OPTIMUM HOTSPOT FINDER APP NOW

optimum

LIKE SEARCHBLOG

Like Send 1,172 people like this. Be the first of your friends.

CATEGORIES

Select Category ▼

MOST ACTIVE POSTS

[Is US Culture Veering Toward The Dark and Deadpan?](#)

8 comments · 1 day ago

[Four Minutes on the Future of Marketing](#)

9 comments · 4 days ago

[Great Content, Meet Great Targeting \(And Reach\)](#)

6 comments · 2 weeks ago

TOP POSTS

[Predictions 2012: The Roundup](#)

[Searchblog 2011: The Year In Writing](#)

[Updated: On FareCast: Rip Me Off No More](#)

[The Database of Intentions](#)

FM'S WHITEBOARD

[Federated Media Launching Native and Programmatic Ad Service Blend - Sales and Marketing @ FolioMag.com](#)
August 23, 2013

["We are trying to make it as easy as possible for people who have thoughtful things to say to get..."](#)
August 23, 2013

[Media Agencies Sign on for BuzzFeed's Course in Creating Buzzy Content](#)
August 22, 2013

[Author Spotlight: Anna Beth Chao of AB Chao FM: Where do you...](#)
August 21, 2013

[Colorful, patterned, eccentric socks are trending at FM amongst...](#)
August 20, 2013

One Question Site Survey
PLEASE ONLY RESPOND TO ANSWER BELOW

How would you rate your opinion of the Ford C-Max?
 SELECT ONE ANSWER

Very Favorably

Favorable

Neutral Opinion

Somewhat Unfavorably

Not at all Favorable



A LITTLE ABOUT JOHN



John Battelle, 45, is an entrepreneur, journalist, professor, and author who has founded or co-founded scores of online, conference, magazine, and other media businesses.
 This work is licensed under a [Creative Commons Attribution - NonCommercial- NoDerivs 2.5 License](#).
[Click to return to my rights page](#) - 2013 John Battelle. | [Terms of Service and Privacy Policies](#)

RELATED LINKS

- [Contact](#)
- [Disclosures](#)
- [Federated Media](#)
- [That Book I Hear He Wrote](#)
- [The Web 2 Conferences](#)

The New York Times
nytimes.com

STEVE CARELL
TONI COLLETTE

September 15, 2004

Amazon to Take Searches on Web to a New Depth

By JOHN MARKOFF

PALO ALTO, Calif., Sept. 14 - Amazon.com, the e-commerce giant, plans to take aim at the Internet search king [Google](http://Google.com) with an advanced technology that the company says will take searches beyond mere retrieval of Web pages to let users more fully manage the information they find.

A9.com, a start-up owned by Amazon, said in a briefing here on Tuesday that it planned to make the new version of its search service, named A9.com, available Tuesday evening. The service will offer users the ability to store and edit bookmarks on an A9.com central server computer, keep track of each link clicked on previous visits to a Web page, and even make personal "diary" notes on those pages for viewing on subsequent visits.

"In a sense, this is a search engine with memory," said Udi Manber, a computer scientist who was a pioneer in online information retrieval and worked at [Yahoo](http://Yahoo.com) before moving to Amazon two years ago.

Mr. Manber created the original A9 search service, which is based in part on search results from Google. He also led the development of Amazon's "search inside the book" project, which lets visitors to the Amazon.com and A9.com Web sites search the complete contents of more than 100,000 books the company has digitally scanned.

Amazon's entry into the search engine wars will certainly raise the stakes in an already heated battle for control of what is believed to be the high ground in Internet commerce and advertising.

Google, which had a widely watched public stock offering last month, is still the dominant provider of search results with approximately 250 million daily searches. But Yahoo and [Microsoft](http://Microsoft.com) have become direct competitors, and a number of start-up companies are busy developing search technologies.

Google executives did not return calls asking for comment.

Amazon is also offering a dialog box that will enable customers on the Amazon.com shopping site to use A9 service to perform Web searches. Company executives say they have no immediate plans to compete head-on with Google and the other search providers. But analysts say the company is aware that search engines are often the starting point for online shopping and cannot help but see broader business opportunities for expanding more fully into online searching.

"They've downplayed the idea that they're going into search," said Danny Sullivan, editor of Search Engine Watch, an industry Web site. "They say, 'we're not competing.' But at the same time you have to wonder why they're doing it, and it's likely they're doing it because they see some potential in search."

Amazon quietly established A9 last year as a subsidiary in a large office building here. The start-up has been offering a search demonstration page, which has so far been limited to the ability to record a history of Web searches.

The new service goes much further, adding the ability to organize and retrieve past searches. The idea is to make searching more useful by making it easier to remember where a Web browser has gone before.

"The ability to search through your own history of personal Web searches is insanely powerful," said John Battelle, a writer and consultant who is the organizer of the Web 2.0 conference to be in San Francisco next month. "This is a big deal," Mr. Battelle said. "But the question is will people get the habit of using it?"

The new A9 search page permits users to search the Web and simultaneously retrieve related information from Google's search results and its image search service, reference material from the GuruNet service and additional information from the Internet Movie Database.

A9 executives said that the new version of the service was simply a first release and that the company had extensive plans for adding new capabilities.

"This is just version 1.0," said Mr. Manber. "There is a lot more to come."

But Mr. Manber, who began working on information retrieval in the early 1990's as a faculty member at the University of Arizona, was reticent to discuss whether A9 would become a direct competitor to Google.

A9 is currently using Google search results and displaying the syndicated Google Adwords advertisements. The two companies share revenue from the advertisements. Amazon also has its own independent technology for indexing the Web, as a result of its purchase in 1999 of Alexa, a search company founded by the information retrieval specialist Brewster Kahle. The new version of A9 offers some Web traffic information derived from Alexa, but not search results.

Initially, A9 will focus on managing information like bookmarks and search history, Mr. Manber said. "It's not just about search," he said. "It's about managing your information."

The A9 service will include a Web browser tool bar that has several innovative features, like the ability to create instant lists from individual Web pages and then use the lists to move among those pages.

Moreover, it will offer a home page giving users the ability to edit and move Web links easily for later retrieval.

The A9 site will also offer a "discovery" feature that gives Internet browsers suggestions on Web sites that they may find interesting, based on their searches - a feature similar to the product recommendation features offered on Amazon.

Mr. Manber said that A9 had no current plans to include paid ads in search research or to give a preference to products sold on Amazon. But he also said that he could not comment on future plans, except to say that A9 did have plans for new search technologies that would generate revenue.

He stressed that the evolution of Internet search capabilities was still in its earliest stages. "We're in the Wright brothers phase of search technology," he said.

A9 executives said they were acutely aware of potential privacy concerns raised by the personalized nature of the service and said they were doing a variety of things to address the issue.

There will be a version of the A9 service that will offer anonymous searches, for example, Mr. Manber said. Moreover, it will be possible to turn off the history feature, remove information from an individual history list and even entirely clear the history results that are stored on the A9 server, he said.

"The new thing here is not that this information is being collected," he said, but rather that A9 is actually letting Web users have access to their browsing histories for their own purposes.

ClickZ (<http://www.clickz.com>)ClickZ Intel (<http://intel.clickzacademy.com>)SES Conference (<http://sesconference.com>)

Search Engine Watch

(<http://searchenginewatch.com/>)



(<http://searchenginewatch.com/author/1910/chris-sherman>)

Amazon's A9 Launches Visual Yellow Pages

Chris Sherman (<http://searchenginewatch.com/author/1910/chris-sherman>), January 26, 2005

• [Page One Power](http://www.pageonepower.com/?utm_source=industry+page&utm_medium=cpc&utm_campaign=SEW+industry+page+) is a custom link building company located in Boise, Idaho. (http://www.pageonepower.com/?utm_source=industry+page&utm_medium=cpc&utm_campaign=SEW+industry+page+)

A9's (<http://a9.com/>) new nationwide yellow pages directory comes with a twist—thumbnail images of business storefronts that let you take a virtual walk through the streets of 10 major U.S. cities.

A9's new offering works much like other online yellow page directories. You can search by business name or category, and get a list of results that are near your current location, with each identified with an icon on a map. Click on a listing for a business, though, and the results are very different than what you see with other services.

In addition to basic business information, you also see a photograph of the business storefront. Beneath this is a series of thumbnail images showing other buildings on either side of the business. Click any of these thumbnails and the main image changes as well. By clicking sequential images you're effectively "moving" along the street, seeing businesses just as you would if you were walking down the street.

"We allow people from their computer to look at the street, to walk to the left, to walk to the right, to see the neighborhood, to see parking—it's virtually like you're there," said Udi Manber, president of A9.com. "Pictures get you information faster than any other way. Very often you remember a place but not its name. This is a very easy way to find it."

To create this system, A9 developed technology that combined digital cameras, GPS location systems and other hardware, and equipped vehicles with this technology to capture images of buildings and geotag them with location information. The company then hired drivers to cover "tens of thousands of miles" to acquire images, according to Manber.

The system is almost entirely automated, and drivers make their rounds through cities at normal speed. "We can do that at a very large scale very efficiently and at very low cost," said Manber.

The system is similar to the building facade photographs available through France Telecom's Yellow Pages site, Pages Jaunes (<http://photos.pagesjaunes.fr/>) that I wrote about last year (<http://searchenginewatch.com/searchday/article.php/3309641>). While it took professional photographers years to assemble the images used in the Pages Jaunes directory, it took A9 only a "few weeks" to capture the 20 million plus images used in its yellow pages system, according to Manber.

A9's yellow pages also lists businesses immediately adjacent to the one you're viewing, as well as those on the opposite side of the street, showing what's located on a full city block. While similar to Metrobot's (<http://searchenginewatch.com/searchday/article.php/3298971>) grid-based search results, it's sparser, allowing you to see at a glance all of the businesses nearby.

A9 will allow business owners to modify listings and send additional content for inclusion in business listings, such as interior photographs, links to web sites and so on, for no charge.

To do so, use the Update Business Info button on the right-hand side of your listing in the Yellow Pages. At the moment, anyone can claim to be the business owner -- actual verification looks to be something that might come over time. The A9 help pages (<http://www.amazon.com/gp/yp/bizedit/help-bizedit.html/>) explain more

This is a clear show over the bow of Yahoo, which rolled out enhanced business listings (<http://searchenginewatch.com/searchday/article.php/3444861>) for Yahoo Local last year. Yahoo charges \$9.95 per month for enhanced listings.

A particularly cool feature allows you to call a business simply by clicking a button on a web page. Enter your phone number, click the button, and the system will automatically initiate a call between your phone and the business' phone.

Want to use the new service? The best way is to go to Amazon itself, where it has become a tabbed service. From the new Yellow Pages (http://www.amazon.com/gp/browse.html/ref=sd_str_yp/?node=3999141) area, you can enter a location in the main box and what you're searching for in the search box on the left-hand side of the screen, to get results.

Oddly, there's no button for it if you go to the A9 home page (<http://a9.com/>) on the right-hand side, where other buttons for History, Bookmarks, Discover and Diary are located.

Instead, over there, you have to perform a search -- say for dentists (<http://a9.com/dentists>) -- then select the Yellow Pages button that appears with others. Now you can see Yellow Pages results, though to localize them, you have to enter an address, city and state or US zip code.

Alternatively, you can bookmark (<http://a9.com/the?a=oyp>) this page. That generates an all-Yellow Pages results page matching the word "the." Change to the query you want to do, change from the default location that appears to what you want, and localized results will be returned.

Hopefully, we'll see a better interface into the system emerge at A9. The ability to enter a location and terms before conducting a search -- and to conduct only a Yellow Pages search -- would be welcome. Even going to the advanced search page (<http://a9.com/-/search/advSearch>), nothing like this is currently available.

Want to learn more about the service direct from A9? A How We Did It (<http://a9.com/-/company/YellowPages.jsp>) page provides video clips of the mapping, plus a map of cities covered.

Like other Amazon services, A9 yellow pages listings also include reviews, ratings, lists and other user-contributed content. Manber says that at some point the system will also add recommendations to search results, similar to the book, music and other recommendations that you get at Amazon.com.

A9 plans to aggressively continue its photography of U.S. cities, adding photographs on virtually a continual basis going forward. A number of cities beyond those included at launch have already been photographed and will go live on the service very soon, once quality assurance has been completed.

A9's entry into the online yellow pages space is sure to quicken both the pace of development and the competitive jostling among all of the players in both online yellow pages and local search listings. For something that was rolled out so quickly, it's already an impressive and useful service that is bound to improve rapidly. Manber declined to talk about any specific future developments, but acknowledged that continued development of the online yellow pages was a high priority for A9.

It's increasingly looking like 2005 might finally be the year when local search grows up and fulfills the promise that's been talked about for years.

Want to discuss or comment on this story? Join the A9 Joins Online Yellow Pages Fray (<http://forums.searchenginewatch.com/forum/showthread.php?t=3944>) discussion in the Search Engine Watch forums.

Our weekly roundup of search engine forum postings will appear in Monday's issue of SearchDay.

Search Headlines

NOTE: Article links often change. In case of a bad link, use the publication's search facility, which most have, and search for the headline.

Registering Keywords... (http://c.moreover.com/click/here.pl?1267919712&f=3000000280910) High Rankings (http://www.highrankings.com/issue128.htm#seo) () Jan 27 2005 11:48PM GMT
Google AdWords API Launched... (http://c.moreover.com/click/here.pl?1267917131&f=3000000280910) Search Engine Watch (http://blog.searchenginewatch.com/blog/050127-183506) () Jan 27 2005 11:43PM GMT
Personalized Search... (http://c.moreover.com/click/here.pl?1267909556&f=3000000280910) IMedia Connection (http://www.imediainconnection.com/content/4977.asp) () Jan 27 2005 11:31PM GMT
Report: Local Online Ad Spend to Rise 46 Percent in '05... (http://c.moreover.com/click/here.pl?1267830847&f=3000000280910) ClickZ Today (http://www.clickz.com) () Jan 27 2005 9:22PM GMT
So what are you reading these days?... (http://c.moreover.com/click/here.pl?1267652175&f=3000000280910) Technology Review (http://www.technologyreview.com/articles/05/02/issue/forward_content.asp?p=1) () Jan 27 2005 4:46PM GMT
Sponsored Links Extend Their Reach... (http://c.moreover.com/click/here.pl?1267587520&f=3000000280910) eWeek (http://www.eweek.com/article2/0,1759,1755439,00.asp?kc=EWRSS03119TX1K0000594) () Jan 27 2005 3:09PM GMT
Custom Tailor a Web Browser Just for You... (http://c.moreover.com/click/here.pl?1267586457&f=3000000280910) New York Times (http://www.nytimes.com/2005/01/27/technology/circuits/27basi.html) () Jan 27 2005 3:08PM GMT
Search engine that speaks... (http://c.moreover.com/click/here.pl?1267586230&f=3000000280910) Sydney Morning Herald (http://www.smh.com.au/news/Breaking/Search-engine-that-speaks/2005/01/27/1106415712128.html) () Jan 27 2005 3:07PM GMT
Translating Arabic for the Web... (http://c.moreover.com/click/here.pl?1267584993&f=3000000280910) Pittsburgh Post-Gazette (http://www.post-gazette.com/pg/05027/448523.stm) () Jan 27 2005 3:05PM GMT
Shutterfly, Yahoo best for sharing photos... (http://c.moreover.com/click/here.pl?1267582878&f=3000000280910) Bluefield Daily Telegraph (http://www.bdtonline.com/articles/2005/01/27/business/02photos.txt) () Jan 27 2005 3:02PM GMT
Yahoo embracing Hollywood scene... (http://c.moreover.com/click/here.pl?1267582088&f=3000000280910) Seattle Times (http://seattletimes.nwsource.com/html/businesstechnology/2002161897_yahoo27.html) () Jan 27 2005 3:01PM GMT
Pay-per-click ads pushing the right buttons... (http://c.moreover.com/click/here.pl?1267580838&f=3000000280910) Globe and Mail (http://www.theglobeandmail.com/servlet/ArticleNews/TPStory/LAC/20050127/TWPPCLIMA27/Business/ldx) () Jan 27 2005 2:59PM GMT
The Yahoo Factor... (http://c.moreover.com/click/here.pl?1267580068&f=3000000280910) Technology Review (http://www.technologyreview.com/articles/05/01/wo/wo_dollarhide012705.asp?lrk=top) () Jan 27 2005 2:58PM GMT



Bringing Together Paid,

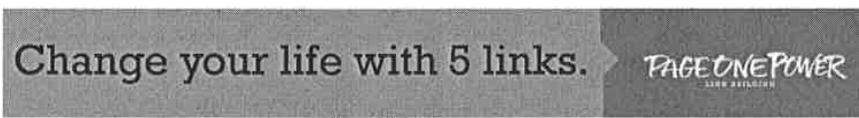
Owned and Earned Media (http://sesconference.com/sanfrancisco/?utm_source=sew&utm_medium=article-promo&utm_content=paid-owned-earned&utm_campaign=sew)

Sept. 10-13, 2013: With a newly announced, completely renovated agenda (http://sesconference.com/sanfrancisco/agenda.php?utm_source=sew&utm_medium=article-promo&utm_content=agenda&utm_campaign=sew),

SES San Francisco could be the most valuable online marketing conference you attend this year. Register today and save up to \$200! (http://sesconference.com/sanfrancisco/registration-details.html?utm_source=sew&utm_medium=article-promo&utm_content=register-save&utm_campaign=sew)

*Pre-show rate through September 6.

()



Incisive Media

(<http://www.incisivemedia.com/>)

Incisive Interactive Marketing LLC.

© 2013 All rights reserved. 55 Broad St, 22nd Floor, New York, NY 10004

Accreditations:



Digital Publisher of the Year 2010 & 2013 (<http://www.incisivemedia.com/accreditations>)



Facebook | Twitter | Breaking news alerts | Register | Sign In

Monday, August 26, 2013

66°F Seattle, WA Mostly Cloudy



Search seattlepi.com Web Search by YAHOO! Businesses

Home Local U.S./World Business Sports Entertainment Life Comics Photos Blogs Forums Education Travel Real Estate Index

Boeing/Aerospace | Microsoft/Tech | Business Wire | Tech Wire | Press Releases

Google hires Amazon's search chief Move deals setback to online rival's ambitions

P-I STAFF AND NEWS SERVICES
Published 10:00 pm, Tuesday, February 7, 2006

0 0 0

Tweet Like Share

Larger | Smaller Email This

Printable Version Font

In a blow to Amazon.com Inc.'s efforts to make a name for itself in search, the head of its online search effort is jumping ship to work for Google Inc.

Udi Manber, who has been chief executive of Amazon's A9 subsidiary, will be a vice president of engineering at Google, spokeswoman Lynn Fox said Tuesday. She declined to provide more details, including when he will start and what specifically he will work on.

Effective immediately, Manber is being replaced at A9 by David Tennenhouse, previously a vice president and director of research at Intel Corp. Amazon spokesman Craig Berman said Manber would be at A9 until Friday.

"It could set back their efforts, but A9 was more of a trial to hedge their bets and have a presence in there," said Safa Rashtchy, a senior Internet analyst for Piper Jaffray & Co. who estimates that 10 percent of Amazon's tech spending centers on search. "I don't think they will give up on it that easily, but it hasn't produced any major advantages for Amazon and it is still experimental."

Google is by far the leader in the search-engine field, commanding 46.3 percent of the U.S. market according to the latest November data from Nielsen/NetRatings. By comparison, Amazon's A9 ranked 29th, with 0.1 percent of the market.

The effect that Google's feat will have upon Amazon's future success in search is unknown. So far, the success of the Mountain View, Calif.-based Google has already pressured many e-commerce companies to expand their technology investments, and Amazon is no exception.

Amazon's main tech investments are on platform, search, Web services and digital, but it is especially hard to discern where the secretive Seattle company's tech spending is focused: on playing catch-up or innovation.

In an e-commerce report, Merrill Lynch analyst Justin Post wrote that he expects Amazon technology investment to reach \$601 million in 2006, a 33 percent jump that is "well ahead of our estimated revenue growth of 20 percent, but a deceleration from 2005 spending growth of 62 percent."

Amazon's spending on technology, product development, and research and development has risen from \$208 million in 2003 to \$406 million in 2005, according to Post's report.

Whether those investments will turn into profits over time remains to be seen, as does what Google expects from its future performance: Unlike most tech companies, Google refuses to provide earnings estimates for financial analysts.

That makes it harder for them to chart the company's growth -- which led to the pummeling of the company's stock last Thursday when it failed to meet their admittedly high expectations.

Manber's defection is the latest coup for Google, a formidable foe for many technology industry titans. In September, the company snagged Internet pioneer Vinton Cerf from MCI Inc. Also last



Ads by Google

Get Your Website For Free

Google Gives You Free Domain and Website. Bring Your Business Online

GYBO.com/Pennsylvania



You Might Also Like

The 12 Worst Supermarkets in America
The Fiscal Times

A Better Way out of Financial Difficulty
BankThink

Sorry, Ben Affleck! 5 Better Choices for Batman
E! Online

fall, Microsoft Corp. and Google waged a court battle over Microsoft executive Kai-Fu Lee's decision to resign from Microsoft to oversee Google's efforts to open a research center in China. That case has since been settled.

Google's other recent notable hires including longtime Microsoft executive Mark Lucovsky and former eBay Inc. executive Louis Monier, who was also a key player in the development of the first search engine, AltaVista.



A billion-dollar daily shot in the arm for the American economy
 ExxonMobil Perspectives Blog

The Death of the PC

The days of paying for costly software upgrades are numbered. The PC will soon be obsolete. And BusinessWeek reports 70% of Americans are already using the technology that will replace it. Merrill Lynch calls it "a \$160 billion tsunami." Computing giants including IBM, Yahoo!, and Amazon are racing to be the first to cash in on this PC-killing revolution. Yet, a small group of little-known companies have a huge head start. Get the full details on these companies, and the technology that is destroying the PC, in a free video from The Motley Fool. Enter your email address below to view this stunning video.

Enter email address... **Click Here, It's Free!** Terms And Conditions Sponsored By The Motley Fool

Printable Version Email This Like Tweet 0 0 Share

Share this story with friends who would like it.

Most Read | Most Commented

1. Woman hits deer, then killed by semi on 405
2. MTV-MVA: What the heck -- Milley Cyrus?
3. Will Smith And His Kids Had The Greatest Reaction Ever To Milley Cyrus' VMA Performance
4. What we learned: Seattle Seahawks at Green Bay Packers (Preseason Week 3)
5. Conn. officer kills lizard found menacing chickens
6. Father: Son missing in woods may be recreating survival film plot
7. How did Sounders fans make that big banner?

You Might Also Like

			
The Amazing Things That GPS Can Do The Boeing Company	Cause Of Yuri Gagarin Death Finally Revealed By Fellow Cosmonaut redOrbit	The 6 Worst Cars at the Detroit Auto Show 2013 The Fiscal Times	Stunning Destinations: San Francisco's Tidepools Expedia Viewfinder

From Around the Web

- 4 Stupid Investment Mistakes Nearly Everyone Makes (etrade.com)
- 16 Countries Where People Work the Longest Hours (The Fiscal Times)
- Why Californians Paid 80 Cents More Per Gallon of Gas (ExxonMobil Perspectives Blog)
- How to Invest Your 401k Savings (WSJ.com: US)
- Gas price economics: What am I paying for in the price of gasoline (ExxonMobil Perspectives Blog)
- Britney Spears' Skin-Tight Church Outfit Would Make God Blush (CafeMom)

We Recommend

- Will Airbus Turn Boeing's Dreamliner Into a Nightmare?
- Can Ford's CEO Save Microsoft?
- Microsoft says CEO Ballmer to retire in 12 months
- New York Times website goes down
- China says economy stabilizing after long slowdown
- Corn prices drop 4 percent; metals, oil rise

Recommended by

Comments

 Add a comment...

Post to Facebook Posting as Michael Boni (Not you?) **Comment**

Facebook social plugin

Reader Service | REPLAY | feeds | Follow us on Twitter | Facebook | E-mail newsletters | Corrections
 Company Information | Send us tips | Job openings | About the P-I | Hearst Corp. | Terms of use | Privacy policy | Your California Privacy Rights | PPC Management by Metrix4Media | Ad Choices

Send comments to newmedia@seattlepi.com

© 2013 Hearst Communications Inc.
HEARST newspapers

6
 6PM.COM™

ALWAYS
 FREE SHIPPING
 SHOP NOW 

REPLAY

HSS
 MY Yahoo!



News from Google

[Subscribe](#)

Google Announces New Mapping Innovations at Where 2.0 Conference

Announcement

May 29, 2007

This morning at the Where 2.0 conference in San Jose, John Hanke, Director of Google Earth & Maps, announced new innovations for Google Maps that offer a whole new perspective on search: Street View and Mapplets. Available on Google Maps at maps.google.com, Street View and Mapplets further Google's commitment to provide users with the most innovative maps available online and developers with new tools for creating and sharing geographic content.

Street View is a new feature of Google Maps that enables users to view and navigate within 360 degree street level imagery of various cities in the US. Street View provides users with a rich, immersive browsing experience directly in Google Maps, enabling greater understanding of a specific location or area. Street View imagery will initially be available for maps of the San Francisco Bay Area, New York, Las Vegas, Denver and Miami, and will soon expand to other metropolitan areas. By clicking on the "Street View" button in Google Maps, users can navigate street level, panoramic imagery. With Street View users can virtually walk the streets of a city, check out a restaurant before arriving, and even zoom in on bus stops and street signs to make travel plans.

Google also announced the launch of Mapplets, a powerful new tool for developers and consumers alike. Mapplets enables third party developers to create mini applications that can be displayed on Google Maps, much like Google Gadgets are displayed on iGoogle. These Mapplets contain a variety of information, from housing listings to crime data, and tools like distance measurement. Users can select from a wide range of Google and third party Mapplets to display on the Map, essentially creating their own "mashup of mashups" directly on the

Google Maps site, while still enjoying the built-in functionality of Google Maps, such as local search and driving directions. A number of our partners, including WeatherBug, [Booking.com](#) and Platial have already created Mapplets. This feature is available at maps.google.com/preview.

DURIE TANGRI LLP
DARALYN J. DURIE (*Pro Hac Vice*)
ddurie@durietangri.com
JOSEPH C. GRATZ (*Pro Hac Vice*)
jgratz@durietangri.com
DAVID MCGOWAN (*Pro Hac Vice*)
dmcgowan@durietangri.com
GENEVIEVE P. ROSLOFF (*Pro Hac Vice*)
grosloff@durietangri.com
217 Leidesdorff Street
San Francisco, CA 94111
Telephone: 415-362-6666
Facsimile: 415-236-6300

Attorneys for Defendant
Google Inc.

IN THE UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK

THE AUTHORS GUILD, INC., Associational
Plaintiff, BETTY MILES, JOSEPH
GOULDEN, and JIM BOUTON, on behalf of
themselves and all other similarly situated,

Plaintiffs,

v.

GOOGLE INC.,

Defendant.

Civil Action No. 05 CV 8136 (DC)

**DECLARATION OF JOSEPH C. GRATZ IN OPPOSITION TO PLAINTIFFS' MOTION
FOR PARTIAL SUMMARY JUDGMENT**

I, Joseph C. Gratz, hereby declare under penalty of perjury:

1. I am an attorney duly admitted to practice law in the State of California and in this Court. I am a member of Durie Tangri LLP, attorneys for Defendant Google Inc. in the above-captioned civil action. I submit this declaration in opposition to Plaintiffs' Motion for Partial Summary Judgment. I make this declaration based on personal knowledge of the facts and circumstances set forth herein.

2. Each of the web page captures attached hereto was captured on August 19, 2013.

3. Attached hereto as Exhibit 1 is a true and correct copy of a portion of the web page resulting for a search on the Library of Congress catalog at <http://catalog.loc.gov> for the query "500 Pearl Street."

4. Attached hereto as Exhibit 2 is a true and correct copy of a portion of the web page resulting for a search on <https://books.google.com> for the query "500 Pearl Street."

5. Attached hereto as Exhibit 3 is a true and correct copy of a portion of the web page that results from clicking on the result *Alas! What Brought Thee Hither?: The Chinese in New York, 1800-1950* in Exhibit 2. This book is in the Google Books Partner Program, and Google displays full pages at the request of its publisher, Fairleigh Dickinson University Press. Because it is in the Partner Program, more than "snippets" of text are displayed.

6. Attached hereto as Exhibit 4 is a true and correct copy of a portion of the web page resulting for a search on <https://books.google.com> for the query "Hong Kee Kang."

7. Attached hereto as Exhibit 5 is a true and correct copy of a portion of the web page that results from clicking on the result *Chinese America, History and Perspectives* in Exhibit 4.

8. Attached hereto as Exhibit 6 is a true and correct copy of the web page located at

<http://www.amazon.com/gp/offer-listing/B002H9DITW/>, which offers for sale used copies of the book appearing in Exhibit 5.

9. Attached hereto as Exhibit 7 is a true and correct copy of Marc Egnal, *Evolution of the Novel in the United States: The Statistical Evidence*, 37:2 SOC. SCI. HIST. 231 (2013).

10. Attached hereto as Exhibit 8 is a true and correct copy of Jean M. Twenge et al., *Changes in Pronoun Use in American Books and the Rise of Individualism, 1960-2008*, 44 J. CROSS-CULTURAL PSYCH. 406 (2013).

11. Attached hereto as Exhibit 8 is a true and correct copy of the written testimony submitted by Paul Aiken to the Committee on the Judiciary of the United States House of Representatives in connection with a September 10, 2009 hearing titled "Competition and Commerce in Digital Books". It is available at <http://judiciary.house.gov/hearings/pdf/aiken090910.pdf>.

12. Attached as Exhibit 9 is a true and correct copy of excerpted pages of the transcript of the deposition of Paul Aiken taken herein on April 19, 2012.

13. Attached as Exhibit 10 is a true and correct copy of excerpted pages of the transcript of the deposition of Betty Miles taken herein on January 4, 2012.

14. Attached as Exhibit 11 is a true and correct copy of excerpted pages of Plaintiffs' Responses and Objections to Defendant Google Inc.'s First Set of Interrogatories to Plaintiffs The Authors Guild, Inc., Jim Bouton, Joseph Goulden and Betty Miles herein, served on April 27, 2012.

15. Attached as Exhibit 12 is a true and correct copy of excerpted pages of the transcript of the deposition of Jim Bouton taken herein on December 15, 2011.

16. Attached as Exhibit 13 is a true and correct copy of excerpted pages of the

transcript of the deposition of Joseph Goulden taken herein on January 6, 2012.

17. Attached as Exhibit 14 is a true and correct copy of Exhibit 6 to the deposition of Paul Aiken taken herein on April 19, 2012.

18. Attached as Exhibit 15 is a true and correct copy of the transcript of the deposition of Daniel Gervais taken herein on June 12, 2012.

19. Attached as Exhibit 16 is a true and correct copy of the transcript of the deposition of Benjamin G. Edelman taken herein on June 14, 2012.

20. Attached as Exhibit 17 is a true and correct copy of Exhibit 5 to the deposition of Jim Bouton taken herein on December 15, 2011.

21. Attached as Exhibit 18 is a true and correct copy of Exhibit 2 to the deposition of Betty Miles taken herein on January 4, 2012.

22. Attached as Exhibit 19 is a true and correct copy of Exhibit 16 to the deposition of Benjamin G. Edelman taken herein on June 14, 2012.

23. Attached as Exhibit 20 is a true and correct copy of excerpted pages of the transcript of the deposition of Judith A. Chevalier taken herein on June 8, 2012.

I declare under penalty of perjury under the laws of the United States that the foregoing is true and correct.

Executed on this 26th day of August, 2013, at San Francisco, California.

/s/ Joseph C. Gratz
Joseph C. Gratz

A-1344

Case 1:05-cv-08136-DC Document 1075-1 Filed 08/26/13 Page 1 of 2

EXHIBIT 1



LIBRARY OF CONGRESS ONLINE CATALOG



- Help
- Search
- Search History
- Headings List
- Titles List
- Request an Item
- Account Info
- Start Over

Try the new version: [Library of Congress Online Catalog](#)

DATABASE: Library of Congress Online Catalog

Your search found no results *

- Check spelling & eliminate punctuation
- Select [another search type](#)
- Enter fewer search words
- Clear search limits
- Use the [help pages](#)
- [Ask a Librarian](#) for assistance or contact your local library.

*** Please note:** The Library of Congress does not keep a copy of every title ever published.

Basic Search **Guided Search**

Search Text:

Search Type: Keyword (match all words)*

- Title Keyword*
- Author/Creator Keyword
- Subject Keyword
- Name/Title Keyword
- Series/Uniform Title Keyword
-
- Expert Search (use index codes and operators)*
- Title Begins With (omit initial article)*
- Author/Creator Browse (enter last name first)
- Subject Browse
- Call Number Browse (LC Classification No.)
- Call Number Browse (Other Shelving Nos.)
- Number Search (LCCN-ISBN-ISSN)*
- Author/Creator Sorted by Title Browse

[*] indicates search limits available

Scroll down for Search Hints

100 records per page ▼

Begin Search

Clear Search

Set Search Limits

Basic Search Tips

Search Type	Brief Help (select a "search type" for detailed Help)	Help on Limits
Keyword (match all words)	- " " surround exact phrases: "war of the worlds" - ? truncates: entrepr? --> finds entrepreneur, entrepreneurial, etc. - Enter more words to increase relevance of search results ("match all words" means that the Boolean "AND" is automatic)	<i>Search Limits Available!</i>
Title Keyword	- Enter any word from the title in any order - Subtitles, alternate titles, translated titles and the statement of responsibility are included in the search - Enter more words to increase relevance of search results	<i>Search Limits Available!</i>
Author/Creator Keyword Subject Keyword Name/Title Keyword Series/Uniform Title Keyword	<i>Headings Keyword Searches:</i> - Enter words from author/creator, subject, name/title or series title areas of the bibliographic record - Enter more words to increase relevance of search results	<i>Search Limits Not Available!</i>
Expert Search (use index codes and operators)	- " " for exact phrases: "tax reform" - Boolean AND, OR, NOT: (mars OR venus) AND mission - Index codes refine searches: KSFG western --> finds western as subject form/genre term	<i>Search Limits Available!</i>

A-1346

Case 1:05-cv-08136-DC Document 1075-2 Filed 08/26/13 Page 1 of 2

EXHIBIT 2

+You Search Images Maps Play YouTube News Gmail Drive Calendar More -

Case 1:05-cv-08136-DC Document 1075-2 Filed 08/26/13 Page 2 of 2

Google "500 Pearl Street"  

Web Images Maps Shopping **Books** More Search tools 

About 14,300 results (0.24 seconds)

[An Act to Designate the Federal Building Located at 500 Pearl ...](#)



books.google.com/books?id=ATqdGAAACAAJ
United States - 2000 - No preview

[Annual reports of the city departments - Page 557](#)



books.google.com/books?id=_E0MAQAAMAAJ
Cincinnati (Ohio) - 1876 - Read - More editions
Pearl street, between Plum street and Central avenue. 500. Pearl street, at corner of Central avenue (double 3-inch). 501. Pearl street, at corner of Central avenue. 502. Pendleton street, between Woodard and Liberty streets (high pressure).

[U.s. Court Directory - Page 350](#)



books.google.com/books?isbn=0788177001
Ann M. Langley - 1998 - Full view
District Southern District of New York Mailing Address Divisional Offices with Resident Deputy John S. Martin, Jr. 212-805-0228 1620 United States Courthouse 500 Pearl Street New York, NY 10007-1312 Lawrence M. McKenna ...

[Document - Volume 26. Part 1. Issues 1-22 - Page 142](#)



books.google.com/books?id=X2hZAAAAYAAJ
1859 - Read - More editions
03 Mulberry street 53 Mulberry street 160 Leonard street 2249 2429 1044 1078 1750 Gas-fitter 500 Pearl street Merchant 472 Pearl street. Bricklayer 15 City-Hall place. Broker ~-~xi Roosevelt street. Merchant 52 Catharine street. Composer .

[Touring Gotham's Archaeological Past: 8 Self-Guided Walking Tours ... - Page 44](#)



books.google.com/books?isbn=0300137893
Diana diZerega Wall - 2004 - Preview - More editions
500 Pearl Street: The Notorious Five Points Slum Walk a short block west to Pearl Street and turn right. Continue north, passing under the ramps leading to the Brooklyn Bridge. Go one block farther, to the corner where Pearl Street turns to the ...

[Documents of the Assembly of the State of New York - Volume 2 - Page 476](#)



books.google.com/books?id=4DMbAQAIAAJ
New York (State). Legislature. Assembly - 1891 - Read - More editions
... school 60 Cooper street Bchool 260 Lamon street school 470 Mullin street school 500 Pearl street school 130 Total 3,010 With the Cooper street and Pearl street schools completed, the former will accommodate 250, and the latter 130 more, ...

[The Legal Researcher's Desk Reference - Page 60](#)



books.google.com/books?id=1gkXAQAAMAAJ
Arlene L. Eis - 2004 - Snippet view - More editions
Janet Bond Arterton New York, NY 10007-1581 Harold Baer, Jr. 141 Church Street (212) 805-0234 500 Pearl Street New Haven, CT 065 10 New York, NY 10007-1312 (203) 773-2456 Kevin Thomas Duffy (212) 805-0184 U.S. Courthouse, ...

[Alas! what Brought Thee Hither?: The Chinese in New York, 1800-1950 - Page 17](#)



books.google.com/books?isbn=0838637043
Arthur Bonner - 1997 - Preview - More editions
He operated a grocery store on Staten Island and, for many y ?ars, owned a cigar business at 500 Pearl Street under the name of William A. Bong? The third recruit may have been John Ah Woh. At an 1878 meeting of Chinese claiming ...

[Sonia Sotomayor: A Biography - Page 107](#)



books.google.com/books?isbn=0313398410
Meg Greene - 2012 - Preview - More editions
During her first few weeks at her office at 500 Pearl Street, Sotomayor struggled with the mounds of material, including printed guidebooks, videos, and seminars that instructed baby judges on the inner workings of the district court. At times ...

A-1348

Case 1:05-cv-08136-DC Document 1075-3 Filed 08/26/13 Page 1 of 2

EXHIBIT 3

+You Search Images Maps Play YouTube News Gmail Drive Calendar More ▾

Google "500 Pearl Street"

Books

Page 17 ▾ < >

Result 1 of 1 in this book for "500 Pearl Street"

No eBook available

Amazon.com
Barnes&Noble.com
Books-A-Million
IndieBound

Find in a library
All sellers >


★★★★★
0 Reviews

Alas! what Brought Thee Hither?: The Chinese in New York, 1800-1950

By Arthur Bonner

"500 Pearl Street"

About this book



New! Get Textbooks on Google Play

Rent and save from the world's largest eBookstore. Read, highlight, and take notes, across web, tablet, and phone.

My library
My History
Books on Google Play



Pages displayed by permission of Fairleigh Dickinson Univ Press.

21 THE 1863 RIOTS TO PROTEST THE CIVIL WAR DRAFT BECAME ATTACKS ON BLACKS. THE BURNING OF THE COLORED ORPHAN ASYLUM ON FIFTH AVENUE LIVED LONG IN THE CITY'S MEMORY.

sibly three, joined the battle. John Akkomb, who in 1856 welcomed a reporter to his Pearl Street boarding house with a polite "How do, Mosheer," was a steward on the gunboat *Massachusetts* in river engagements in Texas and Louisiana. He was twice wounded, once severely in the chest. He returned to the city and lived with his wife, Kitty, and three children on Cherry Street.⁷

Hong Kee Kang, a former sailor who lived on Cherry Street, enlisted in the Navy in July 1863 and served as a steward on the gunboat *Albatross*, handing out powder for the guns when Admiral Farragut blockaded Mobile. He received an honorable discharge in September 1864 and settled down in New York. He operated a grocery store on Staten Island and, for many years, owned a cigar business at **500 Pearl Street** under the name of William A. Hong.⁸

The third recruit may have been John Ah Woh. At an 1878 meeting of Chinese claiming American citizenship, Ah Woh—a cigar maker living on Baxter Street—showed a document dated 1863. Since immigrants were readily granted citizenship as an inducement to enlist, he may have volunteered for service then.⁹

In the early years, the racism faced by the Chinese in New York was a mild reflection of the greater antipathy toward black slaves. In July 1863, New York was swept by five days of rioting touched off by what were seen as unfair laws regulating the drafting of men for the army. At least 105 people were killed and tens of thou-

22 OTHER RIOTERS LYNCHED A BLACK.

sands of dollars worth of property was destroyed. The mobs lynched a black man, burned the Colored Orphan Asylum, and threatened or attacked brothels and homes of racially mixed couples. When someone persuaded a mob that the Chinese were but a "modification" of blacks, the Chinese in the Fourth Ward were also attacked:

Someone disputed this theory, and quite a family quarrel occurred at once. Several blows were struck, the anti-Chinaman in the end getting the worst of it. Afterwards there

17

was a general attack upon all "American citizens of Chinese origin" and John in a state of terror has thrown himself upon the police for his protection. The representatives of this race in New York are a peculiar people. They seem to have a thorough dislike of white men but a very great regard for white women. Most of them are married to females who are only for a husband not particular about his nativity or appearance. The wives "wear the breeches" and John is as obedient

A-1350

Case 1:05-cv-08136-DC Document 1075-4 Filed 08/26/13 Page 1 of 2

EXHIBIT 4

+You Search Images Maps Play YouTube News Gmail Drive Calendar More

Google "Hong Kee Kang"   [SIGN IN](#)

Web Images Maps Shopping **Books** More Search tools 

3 results (0.10 seconds)

[Alas! what Brought Thee Hither?: The Chinese in New York, 1800-1950 - Page 198](#)

books.google.com/books?isbn=0838637043
 Arthur Bonner - 1997 - Preview - More editions
 See also Tong wars Hodgins, Captain "Big Bill," 146 Holocaust, 135 Homer, Winslow, 37 **Hong Kee Kang** (a/k/a Willam A. Hong), 17, 67 Hong Yen Chang, 124 Home, Eph ("Brudder Bones"), 33 Hotel Arlington, 135 Hotel Claridge, 163 Hotel ...

[Surviving the City: The Chinese Immigrant Experience in New York ... - Page 71](#)

books.google.com/books?isbn=0742508919
 Xinyang Wang - 2001 - Preview - More editions
 Sam You and **Hong Kee Kang** — the Civil War veteran — opened their factories at 495 Chatham and 500 Pearl respectively.38 But, for the Chinese in the laundry industry, proximity to customers must have weighed heavily when they were ...

[Chinese America, History and Perspectives - Page 160](#)

books.google.com/books?isbn=1885864000
 Marlon K. Hom - 1995 - Snippet view - More editions
 ... United States.66 As a cigar manufacturer in New York City in 1904, his letterhead read **Hong Kee Kang**.67 And his applications for a certificate to replace his lost discharge papers in 1910 and 1918 are signed John Hang.68 Notwithstanding ...

[Advanced search](#) [Search Help](#) [Send feedback](#)

[Google Home](#) [Advertising Programs](#) [Business Solutions](#) [Privacy & Terms](#) [About Google](#)

A-1352

Case 1:05-cv-08136-DC Document 1075-5 Filed 08/26/13 Page 1 of 2

EXHIBIT 5

+You Search Images Maps Play YouTube News Gmail Drive Calendar More -

Google "Hong Kee Kang"

Books

No eBook available

[AbeBooks](#)
[Amazon](#)

Find in a library
[All sellers »](#)



New! Get Textbooks on Google Play

Rent and save from the world's largest eBookstore. Read, highlight, and take notes, across web, tablet, and phone.

[Go to Google Play Now »](#)

[My library](#)

[My History](#)

[Books on Google Play](#)

Chinese America, History and Perspectives



Marlon K. Hom

★★★★★

0 Reviews

Chinese Historical Society of America, 1995 - Chinese Americans

From inside the book

1 page matching "Hong Kee Kang" in this book

Page 160

24, 1883, he did not sign his name but made his mark, and his navy records are under the name John Ah Heng.⁶⁵ At the County Court of Richmond on Staten Island twelve years later, he signed William Hang on his declaration of intention to become a citizen of the United States.⁶⁶ As a cigar manufacturer in New York City in 1904, his letterhead read **Hong Kee Kang.⁶⁷ And his applications for a certificate to**

[Where's the rest of this book?](#)

What people are saying - [Write a review](#)

We haven't found any reviews in the usual places.

A-1354

Case 1:05-cv-08136-DC Document 1075-6 Filed 08/26/13 Page 1 of 2

EXHIBIT 6



Your Amazon.com Today's Deals Gift Cards Sell Help

Off to College Daily Lightning Deals

Save now Sponsored by Angel Soft

Shop by Department

Search

Books

Go

Hello. Sign in Your Account

Try Prime

Cart

Wish List

Books Advanced Search New Releases Best Sellers The New York Times® Best Sellers Children's Books Textbooks Sell Your Books Best Books of the Month

amazontrade-in



Trade In Over a Million Items for an Amazon.com Gift Card

Learn more

Advertisement



Chinese America: history and perspectives, 1996. (Paperback)

Be the first to review this product

Return to product information

Have one to sell?

Amazon protects every purchase with an A-to-Z guarantee.

All New Used from \$14.98

Show FREE Super Saver Shipping offers only (0)

Sorted by Price + Shipping

Used 1-2 of 2 offers

Price + Shipping

Condition

Seller Information

Buying Options

\$14.98

+ \$3.99 shipping

Used - Good

crepepaperheart

★★★★★ 94% positive. (17 total ratings)

In Stock.

Domestic shipping rates and return policy.

Add to cart

or

Sign in to turn on 1-Click ordering.

\$14.99

+ \$3.99 shipping

Used - Very Good

Excellent condition, light crease mark on front cover. Otherwise like new.

BuzyB

★★★★★ 98% positive over the past 12 months. (106 total ratings)

In Stock. Ships from CA, United States.

Domestic shipping rates and return policy.

Add to cart

or

Sign in to turn on 1-Click ordering.

A-1356

Case 1:05-cv-08136-DC Document 1075-7 Filed 08/26/13 Page 1 of 25

EXHIBIT 7

Marc Egnal

Evolution of the Novel in the United States

The Statistical Evidence

This article examines the evolution of the novel in the United States using a remarkable new source, the Ngram database. This database, which spans several centuries, draws on the 15 million books that Google has scanned. It allows researchers to look at year-to-year fluctuations in the use of particular words. Using one of the available filters, the article is based on English-language books published in the United States between 1800 and 2008. But making sense of these data requires a framework. That framework is provided by the four periods that emerge from much recent writing on the novel. Four epochs—the sentimental era (1789–1860), the genteel era (1860–1915), the modern era (1915–60), and the postmodern era (1960–)—define the evolution of the novel and, more broadly, changes in American society and values. The article argues that a study of key words drawn from the Ngram database confirms the existence of these periods and deepens our understanding of them.

Ever since the first critical works appeared in the early US Republic, observers have proposed a variety of approaches to understanding the evolution of American literature. New points of view, new novels, and new bodies of evidence have led scholars repeatedly to refine or completely recast their interpretations. A database drawn from the more than 15 million books that Google has scanned offers just such an opportunity. This article argues that this source of information, called the Ngram database, clarifies and deepens the analysis of US novels. But databases never speak for themselves. They must be examined in the context of interpretive frameworks. This article

Social Science History 37:2 (Summer 2013)

DOI 10.1215/01455532-2074429

© 2013 by Social Science History Association

begins by exploring recent overviews of American literature. The article next discusses the nature of the Ngram database. It then suggests how these data shed light on American fiction from 1800 to the present.

The New Synthesis of American Literature

Quietly and without much fanfare, a new synthesis has reshaped analysis of American literature. This overview, which replaces older interpretations, provides a crucial background for understanding the Ngram data.¹ At first glance, the assertion that there is a current synthesis seems questionable. The emphasis in recent overviews has been on uncertainty, conflicting explanations, and multiple viewpoints. In his introduction to the seven-volume *Cambridge History of American Literature*, the editor, Sacvan Bercovitch (1994), underscores the demise of older patterns of analysis. “For this generation of critics and scholars,” he writes, “American literary history is no longer the history of a certain, agreed-upon group of American masterworks. Nor is it any longer based upon a certain, agreed-upon historical perspective on American writing.” Bercovitch (*ibid.*: 2) continues, pointing to the diversity of recent viewpoints, “In our times, in short, the study of American literary history defines itself in the plural, through volatile focal points of a multifaceted scholarly, critical, and pedagogic enterprise.” The single-volume, multiauthored *Columbia History of the American Novel* (Elliott 1991) echoes these cautionary themes. Recent discussions of the “canon”—the works that all scholars should read and teach—also highlight diversity, conflict, and the questioning of older authorities (*ibid.*: ix–xviii; Alberti 1995).

Yet all is not chaos. What most recent overviews have in common is an emphasis on periods, and not just any set but rather four particular epochs that define the evolution of US literature. In the introduction to the *Columbia History of the American Novel*, the editor, Emory Elliott (1991: xiv), notes: “There is clearly a chronological progression in the book with four historically organized sections introduced by a specialist in each period.” The four are “Beginnings to Mid-Nineteenth Century,” “The Late Nineteenth Century,” “The Early Twentieth Century,” and “The Late Twentieth Century” (*ibid.*). With variations in name, these four periods define the material in the *Cambridge History of American Literature* as well. The first era is the least likely to bear a formal name. But the second has been called the genteel era, the third the modern era, and the most recent the postmodern era. The newly

published *Cambridge History of the American Novel* (Cassuto 2011) highlights the same four divisions. It begins with part 1, “Inventing the American Novel,” which takes the story up to the Civil War. Next comes “Realism, Protest, Accommodation,” which carries this overview to World War I. The third section, “Modernism and Beyond,” examines writing through the 1960s. The final section, “Contemporary Formations,” opens appropriately with the essay “Postmodern Novels.”

This fourfold division is more than a useful set of labels. It ties literature to the evolution of American society in profound ways. The four periods track changes in the United States, as the nation metamorphosed from an overwhelmingly rural society (in the first era) to one at least initially dominated by small towns (in the genteel era); then to one where vibrant cities set the tone (in the modern era); and further to a time characterized by megalopolises, sprawling entities often with decaying cores (in the postmodern era). Similarly, this periodization embodies far-reaching shifts in values, beginning with an initial concern for womanly virtue (in the first era); evolving into the elaborate codes that marked genteel America; continuing with a less rigid set of moral guidelines while gender and racial hierarchies remained (in the modern era); and concluding with a period in which hallowed assumptions about race, gender, and authority gradually crumbled (in the postmodern era). The four-period synthesis suggests no judgment about the canon, the value of straight or queer literature, colonial or postcolonial attitudes, or male or female roles. In that sense it is very much in keeping with the recent, more open sensibility. But it does assert that any reading of literature should be set securely in the context of these epochs and the evolution of US society.

This article reinforces the current synthesis. The material presented below makes clear that this fourfold division, with its deep roots in American society and culture, provides a valuable framework for literary analysis.

The Ngram Database

The Ngram database sheds light on those eras. But before any source, and particularly one as far-ranging as this compilation, can be used, there is a need to consider its nature, strengths, and weaknesses. In 2004 engineers at Google began making digital copies of books. It is a task of Brobdingnagian proportions. Since the invention of the printing press, about 129 million books have been published; Google has now digitized over 15 million of

them. The project encompasses, along with English books, works in seven other languages. Most of the publications have appeared since 1800. English-language authors, for example, published only 500,000 works between 1500 and 1800.

From the volumes scanned, covering the years from 1500 to 2008, Google selected 5.2 million books as the foundation of the Ngram database. The researchers chose this group partly based on the quality of the optical character recognition (OCR) results. They also selected works with the most complete metadata, including date, place, and language of publication. According to a tabulation undertaken in 2010, the largest portion (67 percent) of the more than 500 billion words in this compilation are in English, with smaller collections in French (8 percent), Spanish (8 percent), German (7 percent), Russian (7 percent), Chinese (2 percent), and Hebrew (< 1 percent). Italian was added to the database after these numbers were computed (Michel et al. 2010, 2011).

Anyone visiting the site (ngrams.googlelabs.com) to run English-language word searches finds a short and not entirely satisfactory list of possible filters. The one used in this article is American English. It limits the search to English-language books published in the United States. These works, however, can be of many sorts, including legal or technical documents, reprints of foreign books, sermons, political tracts, and American fiction. Each book is considered only once, so obscure texts are weighted as heavily as best sellers. There are some virtues to dealing with this large basket of works, but before that question is tackled, the other filters for English-language books should be noted. The database can also be searched for British English (works published in Great Britain), English Fiction (fictional works in English without regard to place of publication), and English and English One Million (both dealing with the broader selection of English-language books). Readers can also examine the smaller Ngram databases for books in seven other languages.

Word searches generate graphs, such as the 17 presented in this article, but what exactly is measured? The numbers on the left axis are percentages indicating for each year the proportion of all words in the scanned books that match the term selected. So if you search for *the*, a very common word, the results show that in most years about 5 percent of all words (Google calls them “unigrams”) are a match. More typical results are much lower. For example, the use of *sublime* peaks early in the nineteenth century at about

0.003 percent. So of every 100,000 words in books published in the 1810s, 3 of them are *sublime* (Cohen 2010; Google Books Ngram Viewer 2011; Hotz 2010).

Google allows users of the database to conduct a finer-grained analysis, but that sorting must be done manually. For any word queried, Google provides a list of the books that include the term during a particular time period. So for *sublime* in the American English database for the 10 years beginning in January 1843, some 254 books are listed. All are published in the United States, but roughly half are reprints of works first issued elsewhere, chiefly in Britain. Some are literary works, for example, the stories of Edgar Allan Poe and the novels of the Anglo-Irish novelist Maria Edgeworth. Many of the volumes are reviews from the United States, such as the *Knickerbocker Review* and the *United States Democratic Review*, or reprints of British publications, such as *Punch* and the *Edinburgh Review*. In sum, the Ngram American English database is a useful handbook for what Americans wrote, but it is a better guide for what middle-class Americans were reading.²

This article examines, with an emphasis on novels, the four periods that emerge from recent overviews of American literature. The analysis highlights the changing popularity of various words as revealed in the Ngram American English database. The graphs presented begin in 1800, although Google provides data for earlier years. But the paucity of books published in the United States before the nineteenth century, and particularly the lack of works written by Americans, make the data from the 1780s and 1790s (or earlier) less reliable. For the four periods, I have used the labels and dates that made the best sense to me. Some studies posit the same divisions, but others might move the termini 10 or even 15 years in one direction or another. Since the underlying graph is always presented, the precise dating is less important than the overall trend line.

Sentimental Era, 1789–1860

During the sentimental era the American middle class, which had first emerged as a self-conscious group in the mid-eighteenth century, came to dominate society and shape its values. Like the other epochs, the sentimental era was a period of great change, but its unity came from the celebration of emotions and a determination to encourage new “proper” and pious patterns of behavior. In contrast to the genteel era with its elaborate, settled codes, the

236 Social Science History

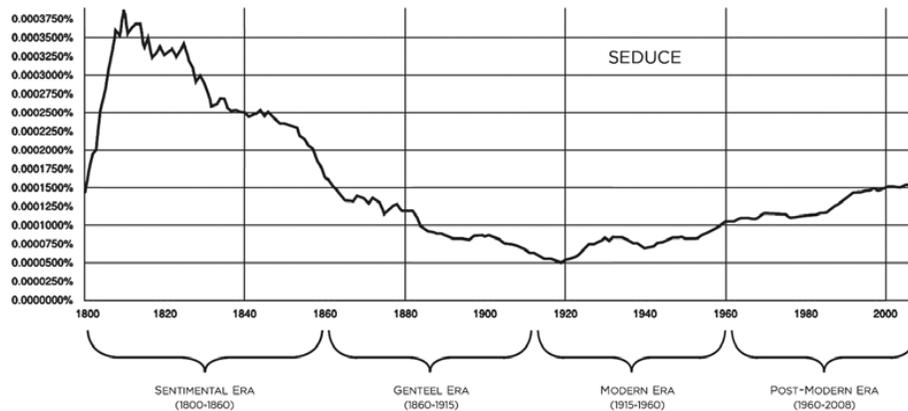


Figure 1 Frequency of *seduce* in American English database

Source: Ngram database.

emphasis in these years was on changing the way people acted (Brown 1940; Davidson 2004; Kete 2000; Tompkins 1985).³

The earliest American novels, such as William Hill Brown's *Power of Sympathy* (1969 [1789]) and Susanna Rowson's *Charlotte Temple* (2011 [1791]), are instructional texts for young women and broadcast the dangers of seduction, particularly by upper-class men. Charles Brockden Brown's novels, such as *Ormond* (1962 [1799]) and *Arthur Mervyn* (1969 [1800]), similarly underscore these perils. While such warnings lessen after 1820, the horrors of seduction and the shame of the "fallen woman" remain evident in such works as Catharine Maria Sedgwick's *Hope Leslie* (1998 [1827]), James Fenimore Cooper's *Deerslayer* (1987 [1841]), and Nathaniel Hawthorne's *Scarlet Letter* (1962 [1850]). Not surprisingly, the literature of this era features the word SEDUCE more often than do later writings. (Capitalized words are displayed in the accompanying graphs.) The word FAITHFUL, a signifier for the proper sort of behavior, also recurs more prominently in this era than in later years (Bercovitch 1994: 620–60; Davidson 2004; Fiedler 1966: 105–23) (figures 1–2).

Efforts to check wanton misdeeds were the opening salvos in a campaign to reshape middle-class values. This concerted effort, which characterized the sentimental era, highlights the growing importance of Christianity and the cult of domesticity. Among the white population, church adherents (those who belonged to a congregation or often attended services) rose from about 14 percent in 1800 to 35 percent in 1850. Not surprisingly, CHURCH

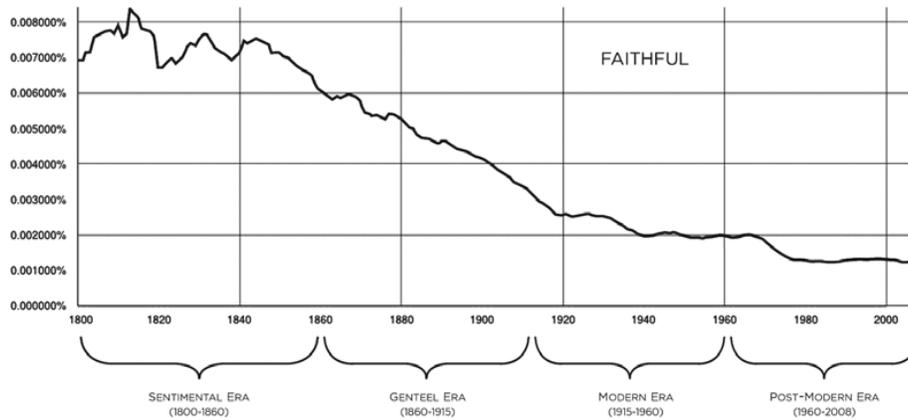


Figure 2 Frequency of *faithful* in American English database
Source: Ngram database.

was widely used during these years—far more so than in later epochs. Similarly characteristic of this first era was the emphasis given to the terms *PIOUS* and *SINFUL* (figures 3–4). The same holds true for the frequency of words (not here graphed) such as *Christianity*, *faith*, *God*, *grace*, and *prayer*. Certainly, the next period, the genteel era, lauded churchgoing—it was part of that epoch’s elaborate codes. However, the heightened concern for piety between 1789 and 1860 reflected the obsessions of a period of intense proselytizing. During these years ministers labored to change people’s beliefs. By contrast, in the genteel era most middle-class individuals knew their duties, and the flood of religious rhetoric receded.⁴

Ann Douglas (1988) describes the changes in this era as the “feminization of American culture” and emphasizes the role of ministers and women authors in this transformation (see also Baym 1993: 11–50; Finke and Stark 1986; Howe 2007: 285–327; Welter 1966). The growing importance of *MOTHER* in comparison to *FATHER* testifies to that feminization and reflects the ascent of the cult of domesticity (figure 5). In 1800 references to “father” are far more common than those to “mother,” but during the ensuing decades, and particularly after 1820, the gap closes strikingly.⁵

Key words highlight two other aspects of the sentimental era. One is the increased appreciation for the natural world. The newfound delight in nature was evident in the work of novelists; in the writings of transcendentalist philosophers, such as Ralph Waldo Emerson; and in the paintings of the Hudson River and Luminist schools. Much of the wonder that Americans felt in

238 Social Science History

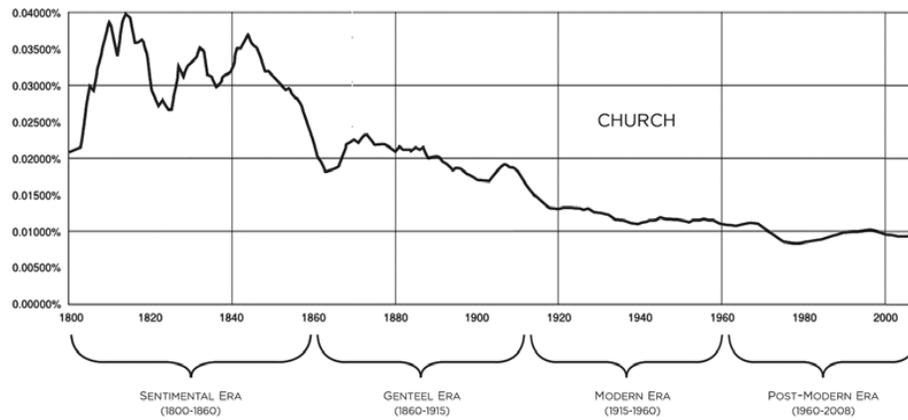


Figure 3 Frequency of *church* in American English database

Source: Ngram database.

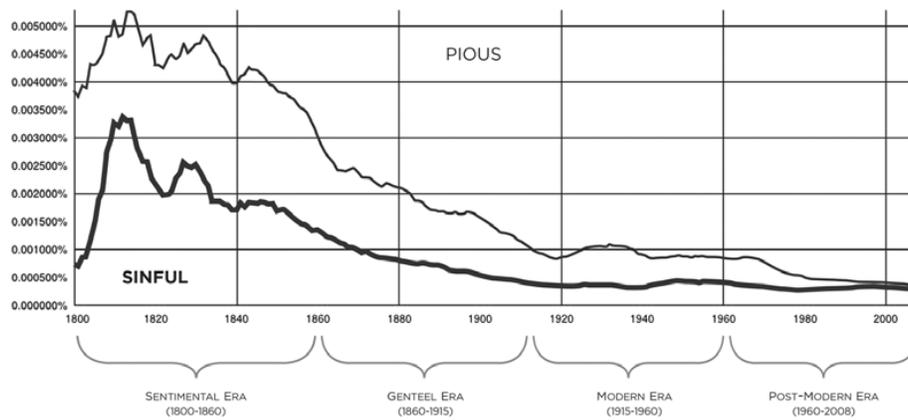


Figure 4 Frequency of *pious* and *sinful* in American English database

Source: Ngram database.

viewing their seemingly limitless forests, streams, waterfalls, and mountains is encapsulated in the word *SUBLIME* (figure 6). The term brings together the qualities of beauty with terror or awe. Many writers, such as Cooper, were fond of this concept. For example, *sublime* and its grammatical variants, such as *sublimity*, appear 20 times in *The Pathfinder* (1961 [1840]) and 8 times in *The Deerslayer* (1987 [1841]).⁶ After 1860 and particularly once industrialization and urbanization had darkened the American landscape, writers were less likely to view the natural world in that exalted light (Blakemore 1986; Novak 2007; Ringe 1971: 3–33).

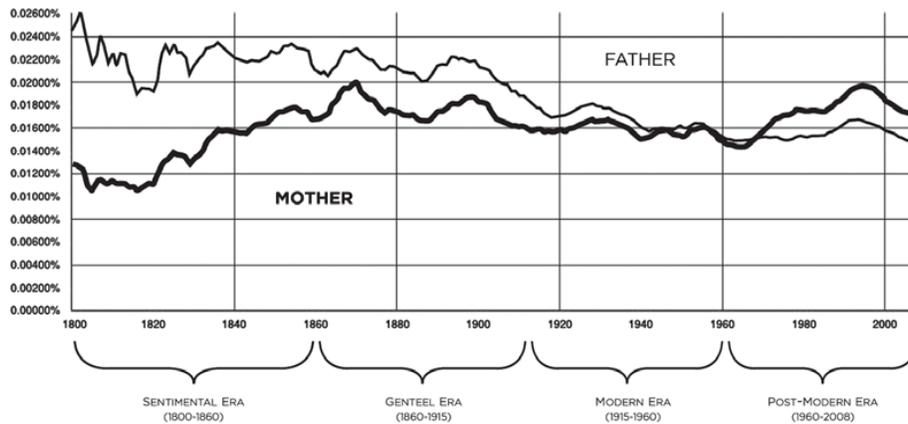


Figure 5 Frequency of *father* and *mother* in American English database

Source: Ngram database.

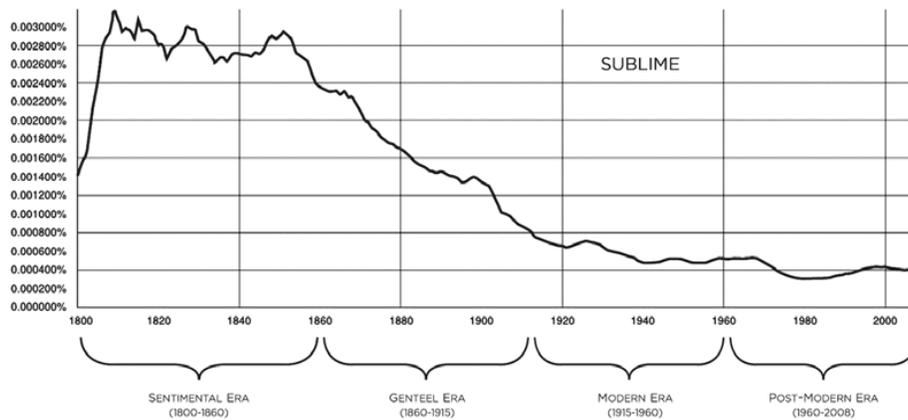


Figure 6 Frequency of *sublime* in American English database

Source: Ngram database.

Still another development important for both literature and, more broadly, American society was the rising importance of the concept of CHILDHOOD (figure 7). The colonists typically had regarded children as miniature adults. During the sentimental era middle-class families gradually came to view these early years as a special stage of development. Beginning in the 1830s, the new outlook was evident in the writings of transcendentalists such as Emerson, the pronouncements of ministers such as Horace Bushnell, and the works of such novelists as Hawthorne, Harriet Beecher Stowe, and Susan Warner. Childhood remained a serious time; not until the genteel era

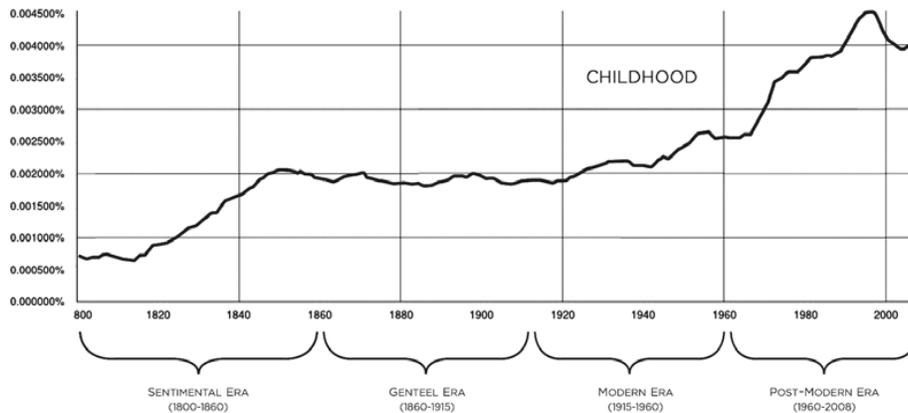


Figure 7 Frequency of *childhood* in American English database

Source: Ngram database.

were adults willing to give children periods of pure play. Still, a new appreciation of the young had emerged (and as the graph suggests, belief in childhood intensified in the postmodern era and will be discussed below) (Mintz 2004; Reinier 1996).

The Genteel Era, 1860–1915

The next epoch, the genteel era, was likewise marked by great changes. What characterized this period was the dominance of an elaborate if unwritten code that defined proper middle-class behavior. This worldview gradually broke down during the last decades of the era, but at its height this credo was marked by a long list of attributes. These included optimism and a clear-cut morality. The genteel code also embraced a belief in the superiority of the white race and a commitment to Protestantism, progress, the doctrine of the two spheres, temperance, and sexual modesty. In 1911, when George Santayana (1967 [1911]) wrote his celebrated essay “The Genteel Tradition in American Philosophy,” many had come to view these beliefs as oppressive. But in its heyday—the 1860s and the 1870s—this outlook encompassed many of small-town America’s noblest virtues, such as an emphasis on caring families and close-knit communities (ibid.; Smith 1978).

Perhaps no word characterizes the goodness of the early genteel era more than *LOVING*, a word widely used from the 1860s to the end of the century (figure 8). The term is the hallmark of Louisa May Alcott’s *Little Women*

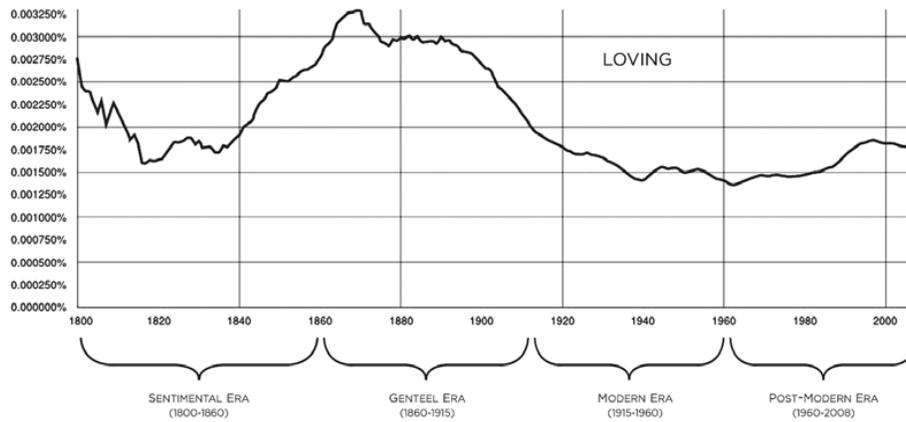


Figure 8 Frequency of *loving* in American English database
Source: Ngram database.

(2008 [1868]) and Mark Twain’s *Tom Sawyer* (1972 [1876]), two novels that epitomize these optimistic decades. Both works are set in small towns. While not everyone is well mannered in Alcott’s *Concord* or Twain’s *Saint Petersburg*, concern for family and neighbors lies at the heart of the books. *Loving* appears 44 times in *Little Women*. When Mr. March goes off to war, he writes to his wife, “I know they [his daughters] will remember all I said to them, that they will be loving children to you . . . that when I come back to them I may be fonder and prouder than ever of my little women.” On another occasion Mrs. March and the girls help a poor German family: “‘That’s loving our neighbor better than ourselves, and I like it,’ said Meg, as they set out their presents while their mother was upstairs collecting clothes for the poor Hummels” (Alcott 2001 [1868]: 52, 58). In *Tom Sawyer*, when the boys return from Jackson’s Island, Huck receives a warm welcome. The narrator comments, “The loving attentions Aunt Polly lavished upon him were the one thing capable of making him more uncomfortable than he was before” (Twain 1972 [1876]: 144).⁷

Other markers of genteel behavior come from the terms of reproach SCOUNDREL and COQUETTE (figure 9). Both denote individuals whose behavior was judged improper but who also remained accepted outliers in Victorian America. In Twain’s *Huckleberry Finn* (2003 [1884]) *scoundrel* is used repeatedly to refer to the two con men, the Duke and the Dauphin. The Duke in fact hurls the epithet at the Dauphin, while Huck refers to both grifters that way. Unlike *scoundrel*, whose use peaked in the genteel

242 Social Science History

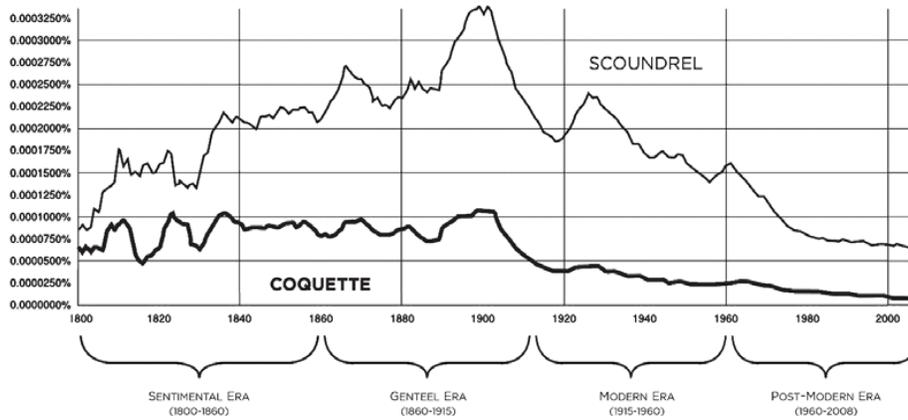


Figure 9 Frequency of *scoundrel* and *coquette* in American English database
Source: Ngram database.

era, *coquette* was widely used in the preceding era as well. Hannah Foster titled her 1797 novel *The Coquette*. Similar to *coquette* was *flirt*, a term whose use soared during the genteel era. In *Little Women* the word appears 17 times both as a noun and as a verb and usually hedged with disapproval. After 1900 the frequency of *scoundrel*, *coquette*, and *flirt* falls off sharply. The framework that cast into sharp relief these mildly transgressive behaviors was itself crumbling.

Another intriguing pair of words, LEG and LIMB, further illustrates the nature of the genteel era and the changes in this period. Until the 1880s the two terms tracked each other, with *limb* serving as a politer referent for *leg*. For example, in *Little Women* Alcott introduces both words. *Limb* is typically used when describing the March sisters. The novel notes Jo's "long limbs" and Beth's "feeble limbs" or how Amy sat "with every limb gracefully composed" (Alcott 2001 [1868]: 48, 241, 312). But when discussing little children, infants, or animals, *leg* becomes the preferred term. Alcott (ibid.: 365, 457) writes, for example, about the "baby's legs" and "Demi's short plaid legs." As the graph makes clear, references to limbs plummet after 1880 (figure 10). This decline, it would seem, is the first indication of the breakdown of the elaborate strictures that came together in the genteel era. More broadly, these changes (and others evident by 1900) suggest an arc that transcends and ties together the four epochs. From the 1780s until the late nineteenth century, American society was marked by an ever more well-defined social code. Beginning at the end of the nineteenth century and continu-

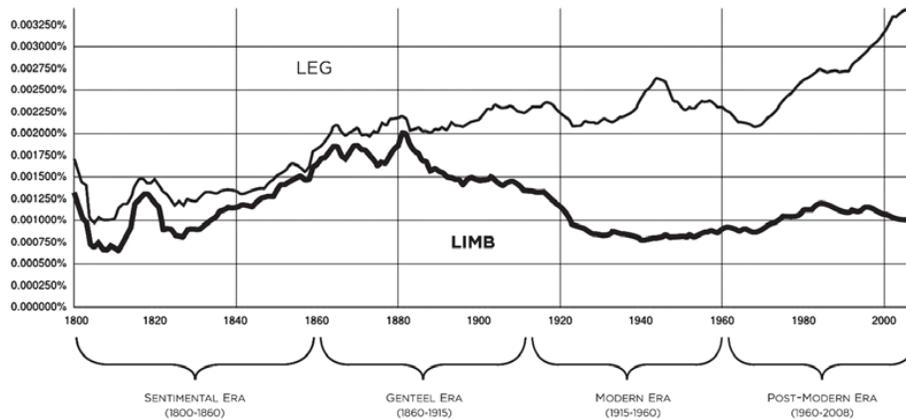


Figure 10 Frequency of *leg* and *limb* in American English database
Source: Ngram database.

ing through the modern and postmodern eras, these prescribed patterns of behavior gradually broke down and were replaced by guidelines that were increasingly less restrictive.⁸

By the turn of the century the genteel world was clearly crumbling, even as the Victorian codes continued to matter. The suicides of the female protagonists in Kate Chopin's *Awakening* (1899) and Edith Wharton's *House of Mirth* (1905) illustrate the force those ideas retained. Still, the decline of older values was unmistakable. Industrialization, urbanization, and the "new" immigration from eastern and southern Europe all undercut the hegemony of small-town, Anglo-Saxon America, which had provided the foundation for the era. A genteel America remained after 1915, lingering particularly in small towns and rural areas, but it no longer dominated the larger society (Delbanco 1997: 83–154; May 1959).

Modern Era, 1915–1960

The changes that sounded the death knell for the genteel era also marked the opening of the modern era. Cities, not small towns, now ruled the social landscape. Many in the middle class no longer deemed sinful those activities, such as drinking or sexual relations outside marriage, that earlier generations had condemned as immoral. Mainstream writers, such as Sinclair Lewis, now reviled rather than praised small towns. But while modern America was in many respects freer and more daring than genteel society had been, it

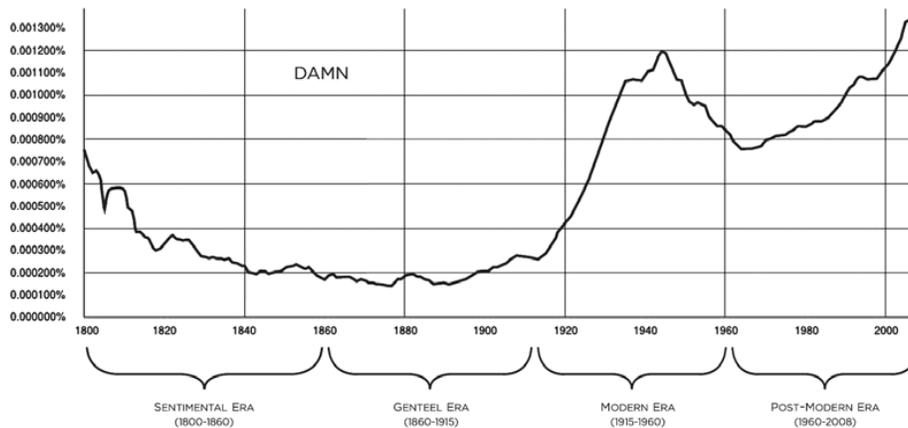


Figure 11 Frequency of *damn* in American English database

Source: Ngram database.

still bore the unmistakable stamp of an established order. Although women now voted and the Harlem Renaissance briefly flourished, gender and racial hierarchies held firm. Few questioned the value of families or religion, even while these institutions were redefined. Even during the Great Depression few doubted that the United States was a wonderful country or that the American dream was worth pursuing (Adams 1978; Baker 1987; Deen 2002; Douglas 1996; Kaplan 2010; Rado 1994).

In part, the modern era was defined by the new freedoms that emerged with the breakdown of genteel values. The decline of *coquette* and *flirt* signaled a new sexual openness. Those who now regarded the “loving” families of Victorian America as stifling institutions celebrated the possibilities that came from franker, more open relationships. The disappearance of *scoundrel* pointed to a society in which individuals undertook grimmer deeds. Few would label as “scoundrels” William Faulkner’s mean-minded protagonists, such as Popeye in *Sanctuary* (1931) or Joe Christmas in *Light in August* (1932). The villains of genteel fiction pale in comparison to the psychologically twisted characters in modern works (Allen 1952; Susman 1984).

The rising use of DAMN is another sign of the loosening of older values (figure 11). This imprecation may seem mild to today’s ears, but the expression, a shortening of *God damn*, was blasphemous and rarely used in literature or public forums before 1915. When Rhett Butler told Scarlett O’Hara in *Gone with the Wind* (1939), “Frankly, my dear, I don’t give a damn,” at least some viewers were shocked (Flamini 1975: 317–20). (The 1930 Motion Pic-

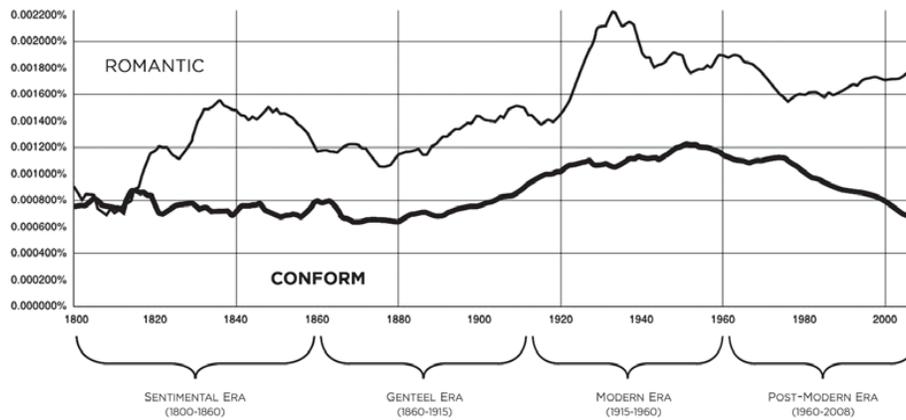


Figure 12 Frequency of *romantic* and *conform* in American English database
Source: Ngram database.

ture Association’s Production Code had banned the word and fined the producer David O. Selznick \$5,000.) The word abounds in modern literature. For example, *damn* or *darned* appears 43 times in F. Scott Fitzgerald’s *This Side of Paradise* (1920) and 35 times in his *Beautiful and the Damned* (1922).

The new sense of freedom that marked the early years of the modern era was also captured in the word *ROMANTIC* (figure 12). Where *loving* suggested the dependable, caring individuals who were the pillars of the Victorian family, *romantic* had a more transgressive feel. It described persons who listened to their own hearts and cast aside society’s codes. *Romantic* had first risen in importance during the heyday of transcendentalism. Its frequency lessened during the early genteel era, when Victorian morality prevailed. But as those strictures weakened, *romantic* became more common, and its use soared in the 1920s. Jay Gatsby in Fitzgerald’s *Great Gatsby* (1925) stands as the quintessential romantic hero. In introducing him, the narrator, Nick Carraway, notes that Gatsby possesses “a romantic readiness such as I have never found in any other person and which it is not likely I shall ever find again” (Fitzgerald 1925: 2).

The modern era had another side, one that was more staid and conformist. The ascent of the term *MIDDLE CLASS* also characterizes this period (figure 13). This designation had been gradually growing in importance since the nineteenth century, but the gains were most striking between 1915 and 1940 (another sharp advance came after 1960). The modern era was an epoch of increasing wealth and comfort despite the hardships of the Great Depres-

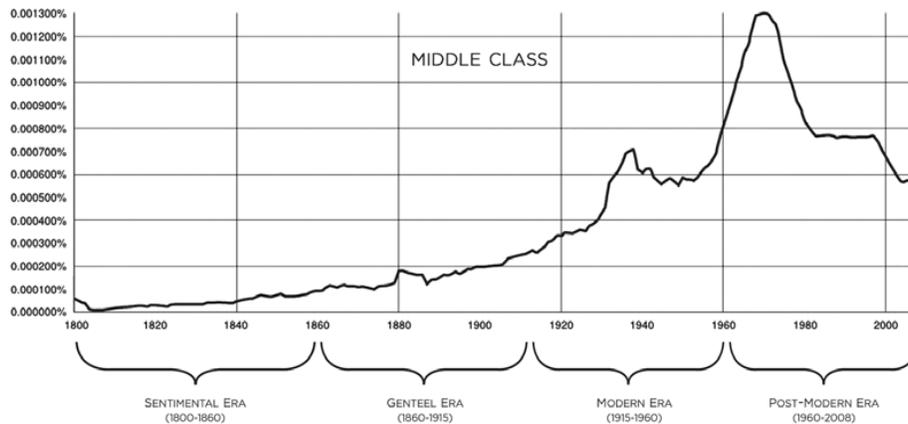


Figure 13 Frequency of *middle class* in American English database
Source: Ngram database.

sion. The Ngram database indicates that the use of words such as *profits*, *earnings*, and *investments* skyrocketed in this era. These terms, to be sure, were found more often in the business press than in fiction, but novels also help detail the expansion of the middle class. Logan Killicks and Jody Starks in Zora Neale Hurston's *Their Eyes Were Watching God* (1937) are exemplars of these new monied individuals, as are the Snopeses in Faulkner's trilogy that begins with *The Hamlet* (1940).

By the 1950s the middle class was prosperous, often smug, and prone to conformity. Significantly, citations of two transgressive terms—*romantic* and *damn*—which had risen in the 1920s and the 1930s, fell off in the 1950s. Instead, the last part of this epoch was marked by the steady ascent of the term CONFORM, whose use peaked in the 1950s (figure 12). In the 1950s novelists such as Ralph Ellison, J. D. Salinger, Mary McCarthy, and Saul Bellow depicted and criticized the middle class. But no writer saw an alternative to accepting and living with this oppressive society. For example, in Salinger's *Catcher in the Rye* (1951: 13) the protagonist, Holden Caulfield, rails against conformist society, complaining that he was “surrounded by phonies,” but in the end plans to become part of that world.

Postmodern Era, 1960–

The postmodern era marks another sea change in American society and culture. Long-held assumptions that had shaped modern society now slowly

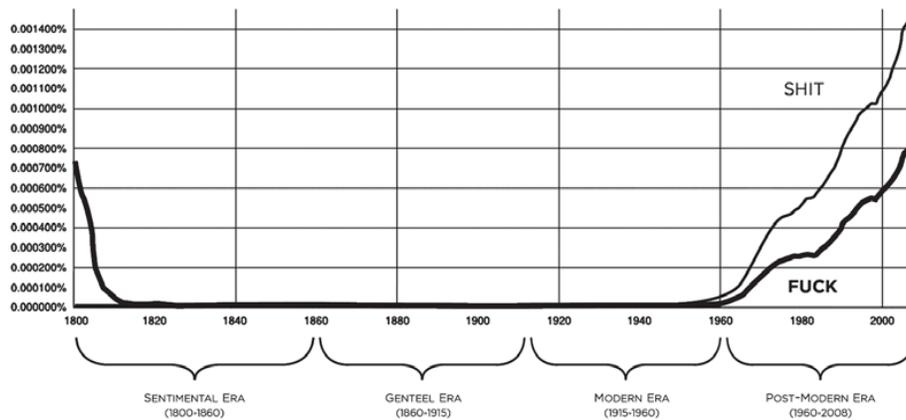


Figure 14 Frequency of *shit* and *fuck* in American English database
Source: Ngram database.

crumbled, much as the edifice of Victorian values had shattered early in the twentieth century. No longer did educated, middle-class Americans accept gender and racial hierarchies that had prevailed for centuries. The civil rights and women's movements had a far-reaching impact. (Terms such as *racism* and *sexism*, little used in earlier years, now soared in importance.)⁹ Novelists and others questioned the meaning and value of middle-class life. Characteristic of the era was the waning of the traditional family and the growing acceptance of alternative lifestyles. At the same time, the hollowing out of American industry eliminated a broad swath of the middle class. Changes in government policies heightened the extremes in wealth. Once vibrant cities now housed the gated rich and the teeming poor (Elias 1999; Jameson 1991, 1998; McHale 1987; Steiner 1999).

The growing disdain for older, middle-class values was evident in the spread of profanity. The prevalence of SHIT, FUCK, and other formerly taboo words testified to the new order (figure 14). When Norman Mailer wrote his first novel, *The Naked and the Dead* (1948), his publishers persuaded him to use *fug* instead of *fuck*. Mailer's postmodern works have no comparable circumspection. *Fuck* and *cocksucker* fill the pages of *The Executioner's Song* (1979). Similar language is present in other works of great seriousness, such as the novels of Philip Roth. (Incidentally, the spike between 1800 and 1815 in the graph for *fuck* reflects a flaw in Google's OCR: the older letter *s* was misread as *f*.)

Another change signaled by language is the growing importance of

248 Social Science History

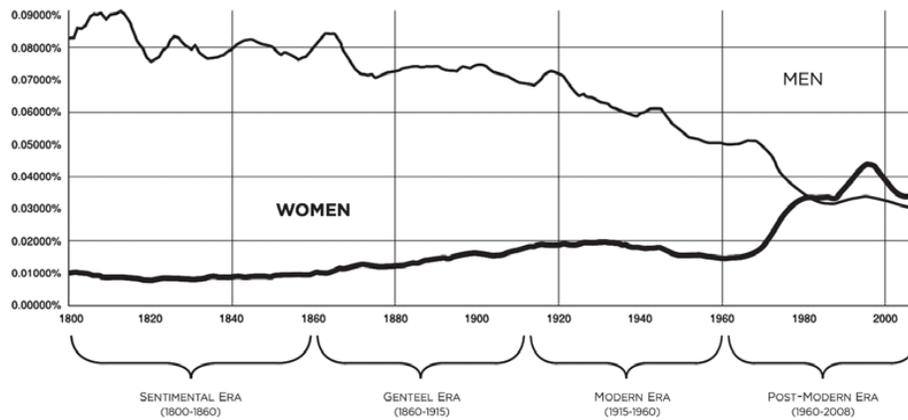


Figure 15 Frequency of *men* and *women* in American English database
Source: Ngram database.

women in society and literature. This trend is evident in the relative frequency of **WOMEN** and **MEN** in printed material (figure 15). During the sentimental era, even as *mother* grew more important, *women* barely rose in frequency. However, since 1960 *men* and *women* have switched places in the frequency of citations. The swap occurred because there were fewer references to “men”—a long-term trend—and because there were many more citations of “women.” *Man* or *men* no longer served as a synecdoche for *people*. Moreover, the use of *mother* has surpassed that of *father* in postmodern books.¹⁰

A similar pattern emerges in looking at **HE** and **I**, with **I** emerging as clearly predominant in the postmodern era (figure 16). References to the two pronouns had long fluctuated together, with **I** holding a slight edge since the 1830s, when transcendentalists and others exalted the individual. In today’s world **I** has become ever more common. The trend is hardly surprising in an era of confessional novels and, more recently, bloggers eager to post their innermost thoughts.

Finally, one of the most striking developments of the postmodern era has been the celebration of ties between parents and children, an emotional bonding reflected in many current novels. To many observers, postmodern society appears chaotic, lacking the structures and rules that once held groups together. In this sometimes frightening new world, the bedrock that remains for many is the love between a parent and a child. The new concern for children is evident in the sharply increased references after 1960 to

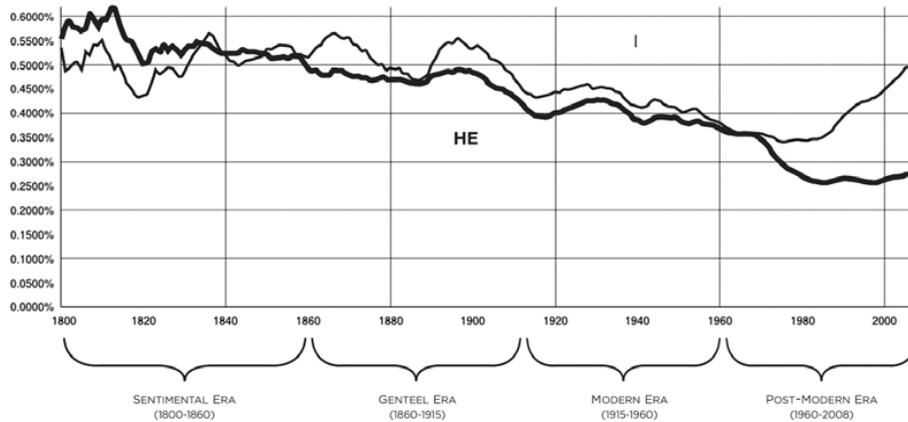


Figure 16 Frequency of *I* and *he* in American English database
Source: Ngram database.

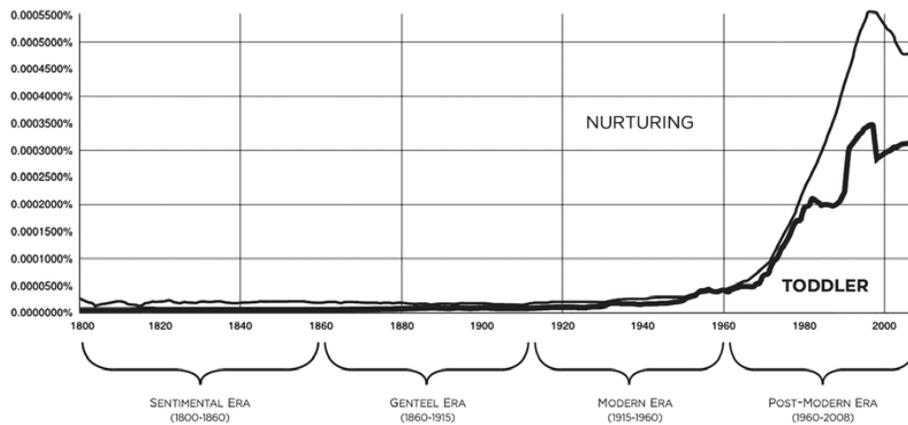


Figure 17 Frequency of *nurturing* and *toddler* in American English database
Source: Ngram database.

NURTURING and TODDLER (figure 17). The trend is also evident in the reinvigoration of the term *childhood*, which first rose in importance during the sentimental era. Many of the key postmodern novels have at their core the tight link between a parent and a child. Works that illustrate this tie are as diverse as Toni Morrison's *Beloved* (1987), Roth's *American Pastoral* (1997), Richard Russo's *Empire Falls* (2001), and Cormac McCarthy's *Road* (2006).¹¹

To conclude, this article argues that the study of various terms drawn from the Ngram database deepens our understanding of the evolution of American novels. Any essay such as this one, which deals with a broad sweep

of literature, a powerful but imperfect database, and links between social change and literature, can hardly be taken as the final word. The conclusions set forth are necessarily hypotheses subject to further testing. Future refinements to the Ngram database or its successors will surely make possible more precise searches.¹² Still these data, which rest on the analysis of millions of books, are too valuable to ignore even in their present form. These data confirm the current division of US literature into four eras and sharpen the conclusions about each of those periods.

Notes

I would like to thank Art Redding, Randolph Roth, Karen Wilson, and Brett Zimmerman for their comments on an earlier version of this article. Ben Egnal created the graphs.

- 1 Typically, earlier overviews sought to unify American literature around a single broad theme, such as the conflict between the “people” and plutocrats (Parrington 1927), the impact of mechanization (Marx 1964), or the differences between English and American novelists (Chase 1957; Fiedler 1966).
- 2 The Ngram database reflects the interests of the middle class more than the outlook of the working class. Google apparently scanned very few of the hundreds of thousands of dime novels that circulated in the nineteenth century and little of the pulp fiction or popular romance literature of the twentieth century. As various scholars have noted, these mass circulation works often present viewpoints at odds with “middlebrow” and “highbrow” literature. For a discussion of this popular literature, see Denning 1987; Goulart 1972; Jones 1978; Reynolds 1989.
- 3 The year 1789, which opens this era, marks the publication of the first American-authored novel, William Hill Brown’s *Power of Sympathy*.
- 4 Other words whose use peaked during this period and which reflect the religious fervor of the sentimental era are *Christ*, *confess*, *devil*, *Lord*, *repent*, and *Satan*. Analysis based on the *Oxford English Dictionary* (Simpson 2012) suggests that the meanings of these words, and indeed of most of the terms examined in the graphs, text, and notes of this article, have not notably changed since the 1780s. For the exceptions to the rule, see the analysis of *leg* and *limb* in the text and the discussions in notes 9 and 11. However, as this article makes clear, the *frequency* with which words have been used has varied dramatically (*ibid.*).
- 5 Other words characteristic of the sentimental era and the newly ascendant cult of domesticity are *affectionate*, *compassion*, *domestic*, and *submissive*.
- 6 E-book editions, available at www.gutenberg.org, made possible the tabulations of Cooper’s word use. This same source was used for the other single-book word counts mentioned in the text.
- 7 An extensive literature examines Twain’s complex relationship to the genteel tradition. See Brooks 1920; Kaplan 1966; Krauth 1999; and Smith 1978: 102–27.

- 8 Various terms whose use declines after 1900 are also markers of the breakdown of genteel values. These words include *brave*, *hero*, *honor*, *house*, *respectful*, *sentiment*, *sigh*, *tender*, and *villain*. *Temperance* also falls off but has a brief second surge in the 1920s with Prohibition.
- 9 According to the *Oxford English Dictionary* (Simpson 2012), *racism* with its modern meaning dates from the 1920s, while *sexism* as currently used emerged only in the 1960s. In earlier eras *sexism* was a rarely used term that referred to belonging to one gender, not discrimination against women (ibid.).
- 10 Similarly, *grandmother* has surpassed *grandfather* in the postmodern era.
- 11 The frequency of other terms that reflect the new devotion to nurturing has soared in the postmodern era. Among these words are *baby*, *babysitter*, *caregiver*, *caring*, *day care*, *infant*, and *nanny*. Two words, *domestic* and *home*, widely used in the sentimental and genteel eras, have experienced a new vogue. The *Oxford English Dictionary* (Simpson 2012) suggests that *babysitter* dates from the 1930s, while only during the nineteenth century did *toddler* (see figure 17) come to mean a small child.
- 12 Readers may wonder whether all the words tested for this article conformed neatly to the four periods. Most did, but a few did not. Some words have multiple meanings. For example, a *rake* can be a cad or a garden implement, and its Ngram pattern shows little clarity. Some words reflect events or particular time periods. *Soldier* peaked during the Civil War and the two world wars; *tenement* soared during the early twentieth century. A few words were simply puzzling. For example, *greedy* and *understanding* fluctuated with little evident pattern.

References

- Adams, Robert Martin (1978) "What was modernism?" *Hudson Review* 21 (1): 19–33.
- Alberti, John, ed. (1995) *The Canon in the Classroom: The Pedagogical Implications of Canon Revision in American Literature*. New York: Garland.
- Alcott, Louisa May (2001 [1868]) *Little Women*. Peterborough, ON: Broadview.
- Allen, Frederick Lewis (1952) *The Big Change: America Transforms Itself, 1900–1950*. New York: Harper.
- Baker, Houston A., Jr. (1987) *Modernism and the Harlem Renaissance*. Chicago: University of Chicago Press.
- Baym, Nina (1993) *Women's Fiction: A Guide to Novels by and about Women in America, 1820–1870*, 2nd ed. Urbana: University of Illinois Press.
- Bercovitch, Sacvan, ed. (1994) *The Cambridge History of American Literature*. Vol. 1, 1590–1820. New York: Cambridge University Press.
- Blakemore, Stephen (1986) "Language and world in *The Pathfinder*." *Modern Language Studies* 16 (3): 237–46.
- Brooks, Van Wyck (1920) *The Ordeal of Mark Twain*. New York: Dutton.
- Brown, Charles Brockden (1962 [1799]) *Ormond*. New York: Hafner.
- (1969 [1800]) *Arthur Mervyn*. Kent, OH: Kent State University Press.

252 Social Science History

- Brown, Herbert R. (1940) *The Sentimental Novel in America, 1789–1860*. Durham, NC: Duke University Press.
- Brown, William Hill (1969 [1789]) *Power of Sympathy*. Columbus: Ohio State University Press.
- Cassuto, Leonard, ed. (2011) *The Cambridge History of the American Novel*. New York: Cambridge University Press.
- Chase, Richard (1957) *The American Novel and Its Tradition*. New York: Doubleday Anchor.
- Chopin, Kate (1899) *The Awakening*. New York: Stone.
- Cohen, Patricia (2010) “In Five Hundred Billion Words, New Window on Culture.” *New York Times*, December 16, www.nytimes.com/2010/12/17/books/17words.html?page-wanted=1&r=1.
- Cooper, James Fenimore (1961 [1840]) *The Pathfinder*. New York: New American Library.
 ——— (1987 [1841]) *The Deerslayer*. New York: Penguin.
- Davidson, Cathy N. (2004) *Revolution and the Word: The Rise of the Novel in America*, expanded ed. New York: Oxford University Press.
- Deen, Stella, ed. (2002) *Challenging Modernism: New Readings in Literature and Culture, 1914–1945*. Burlington, VT: Ashgate.
- Delbanco, Andrew (1997) *Required Reading: Why Our American Classics Matter Now*. New York: Farrar, Straus and Giroux.
- Denning, Michael (1987) *Mechanic Accents: Dime Novels and Working-Class Culture in America*. London: Verso.
- Douglas, Ann (1988) *The Feminization of American Culture*. New York: Doubleday Anchor.
 ——— (1996) *Terrible Honesty: Mongrel Manhattan in the 1920s*. New York: Noonday.
- Elias, Amy J. (1999) “And the beat goes on: Defining postmodernism and its other.” *Contemporary Literature* 40 (1): 161–80.
- Elliott, Emory, ed. (1991) *The Columbia History of the American Novel*. New York: Columbia University Press.
- Faulkner, William (1931) *Sanctuary*. New York: Cape and Smith.
 ——— (1932) *Light in August*. New York: Smith and Haas.
 ——— (1940) *The Hamlet*. New York: Random House.
- Fiedler, Leslie A. (1966) *Love and Death in the American Novel*. New York: Dell.
- Finke, Roger, and Rodney Stark (1986) “Turning pews into people: Estimating nineteenth century church membership.” *Journal for the Scientific Study of Religion* 25 (2): 180–92.
- Fitzgerald, F. Scott (1920) *This Side of Paradise*. New York: Scribner.
 ——— (1922) *The Beautiful and the Damned*. New York: Scribner.
 ——— (1925) *The Great Gatsby*. New York: Scribner.
- Flamini, Roland (1975) *Scarlett, Rhett, and a Cast of Thousands: The Filming of *Gone with the Wind**. New York: Macmillan.
- Foster, Hannah (1986 [1797]) *The Coquette*. New York: Oxford University Press.
- Google Books Ngram Viewer (2011) ngrams.googlelabs.com/info (accessed April 4, 2011).

- Goulart, Ron (1972) *Cheap Thrills: An Informal History of the Pulp Magazines*. New Rochelle, NY: Arlington House.
- Hawthorne, Nathaniel (1962 [1850]) *The Scarlet Letter*. Columbus: Ohio State University Press.
- Hotz, Robert Lee (2010) "World-Wide Web launches: New Google database puts centuries of cultural trends in reach of linguists." *Wall Street Journal*, December 17, online.wsj.com/article/SB10001424052748704073804576023741849922006.html.
- Howe, Daniel Walker (2007) *What Hath God Wrought: The Transformation of America, 1815–1848*. New York: Oxford University Press.
- Hurston, Zora Neale (1937) *Their Eyes Were Watching God*. Philadelphia: Lippincott.
- Jameson, Fredric (1991) *Postmodernism; or, The Cultural Logic of Late Capitalism*. Durham, NC: Duke University Press.
- (1998) *The Cultural Turn: Selected Writings on the Postmodern, 1983–1998*. New York: Verso.
- Jones, Daryl (1978) *The Dime Novel Western*. Bowling Green, KY: Popular Press.
- Kaplan, Carla (2010) "Making it new: Constructions of modernism," in Paul Lauter (ed.) *A Companion to American Literature and Culture*. Malden, MA: Wiley-Blackwell: 40–56.
- Kaplan, Justin (1966) *Mr. Clemens and Mark Twain: A Biography*. New York: Simon and Schuster.
- Kete, Mary Louise (2000) *Sentimental Collaboration: Mourning and Middle-Class Identity in Nineteenth-Century America*. Durham, NC: Duke University Press.
- Krauth, Leland (1999) *Proper Mark Twain*. Athens: University of Georgia Press.
- Mailer, Norman (1948) *The Naked and the Dead*. New York: Rinehart.
- (1979) *The Executioner's Song*. Boston: Little, Brown.
- Marx, Leo (1964) *The Machine in the Garden: Technology and the Pastoral Ideal in America*. New York: Oxford University Press.
- May, Henry F. (1959) *The End of American Innocence: The First Years of Our Own Time, 1912–1917*. New York: Oxford University Press.
- McCarthy, Cormac (2006) *The Road*. New York: Vintage.
- McHale, Brian (1987) *Postmodernist Fiction*. New York: Methuen.
- Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak, and Erez Lieberman Aiden (2010) "Supporting online material for quantitative analysis of culture using millions of digitized books." *Science*, December.
- (2011) "Quantitative analysis of culture using millions of digitized books." *Science*, January 14, 176–82.
- Mintz, Steven (2004) *Huck's Raft: A History of American Childhood*. Cambridge, MA: Harvard University Press.
- Morrison, Toni (1987) *Beloved*. New York: Knopf.
- Novak, Barbara (2007) *Nature and Culture: American Landscape and Painting, 1825–1875*, 3rd ed. New York: Oxford University Press.

254 Social Science History

- Parrington, Vernon Louis (1927) *Main Currents in American Thought: An Interpretation of American Literature from the Beginnings to 1920*. New York: Harcourt, Brace.
- Rado, Lisa, ed. (1994) *Rereading Modernism: New Directions in Feminist Criticism*. New York: Garland.
- Reinier, Jacqueline S. (1996) *From Virtue to Character: American Childhood, 1775–1850*. New York: Twayne.
- Reynolds, David S. (1989) *Beneath the American Renaissance: The Subversive Imagination in the Age of Emerson and Melville*. Cambridge, MA: Harvard University Press.
- Ringe, Donald A. (1971) *The Pictorial Mode: Space and Time in the Art of Bryant, Irving, and Cooper*. Lexington: University of Kentucky Press.
- Roth, Philip (1997) *American Pastoral*. Boston: Houghton Mifflin.
- Rowson, Susanna (2011 [1791]) *Charlotte Temple*. New York: Norton.
- Russo, Richard (2001) *Empire Falls*. New York: Vintage.
- Salinger, J. D. (1951) *The Catcher in the Rye*. Boston: Little, Brown.
- Santayana, George (1967 [1911]) “The genteel tradition in American philosophy,” in Douglas L. Wilson (ed.) *The Genteel Tradition: Nine Essays by George Santayana*. Cambridge, MA: Harvard University Press: 37–64.
- Sedgwick, Catharine Maria (1998 [1827]) *Hope Leslie*. New York: Penguin.
- Simpson, John A., ed. (2012) *Oxford English Dictionary*, www.oed.com (accessed April 10, 2012).
- Smith, Henry Nash (1978) *Democracy and the Novel: Popular Resistance to Classic American Writers*. New York: Oxford University Press.
- Steiner, Wendy (1999) “Postmodern fictions, 1970–1990,” in Sacvan Bercovitch (ed.) *The Cambridge History of American Literature*. Vol. 7, *Prose Writing, 1940–1990*. New York: Cambridge University Press: 427–538.
- Susman, Warren I. (1984) *Culture as History: The Transformation of American Society in the Twentieth Century*. New York: Pantheon.
- Tompkins, Jane (1985) *Sensational Designs: The Cultural Work of American Fiction, 1790–1860*. New York: Oxford University Press.
- Twain, Mark (1972 [1876]) *Tom Sawyer*. New York: Pocket.
- (2003 [1884]) *Huckleberry Finn*. Berkeley: University of California Press.
- Welter, Barbara (1966) “The cult of true womanhood, 1820–1860.” *American Quarterly* 18 (2, pt. 1): 151–74.
- Wharton, Edith (1905) *House of Mirth*. New York: Scribner.

A-1381

Case 1:05-cv-08136-DC Document 1075-8 Filed 08/26/13 Page 1 of 31

EXHIBIT 8

**Statement of Paul Aiken
On the Google Book Settlement**

**Committee on the Judiciary
House of Representatives**

September 10, 2009

I. EXECUTIVE SUMMARY

Overview

My name is Paul Aiken. I'm the executive director of the Authors Guild, the largest society of published authors in the U.S., representing more than 8,500 book authors and freelance writers. Our members represent the broad sweep of American authorship, including literary and genre fiction, nonfiction, trade, academic, and children's book authors, textbook authors, freelance journalists and poets.¹ Guild members have won countless honors and all major literary awards, including the Nobel Prize for Literature.²

The Authors Guild promotes the professional interests of authors: we're advocates for effective copyright protection, fair contracts, and free expression.

The Challenges Facing Print Media

It's a pleasure to be here before this committee, at this moment in book publishing history. Never in the Authors Guild's long history has its straightforward mission – to maintain writing as a viable livelihood – been so daunting. The digital environment is brutal for print media. As we meet here today, the newspaper industry is dying. Credible estimates say that one newspaper is closing each week in America. The magazine industry isn't much better off, as week by week we see venerable publications shrink in size and ambition. The loss to our society from the collapse of these industries is

¹ The Guild had its beginnings as the Authors League of America, which was founded in 1912 by a group of book authors (including Theodore Roosevelt, who served as the League's founding vice president), short story writers, freelance journalists and a smattering of dramatists. In the 1920s, the Authors League broke into two groups: the Authors Guild and the Dramatists Guild of America.

² Pearl S. Buck (1938) (who served as Authors Guild president), William Faulkner (1949), John Steinbeck (1962), and Isaac Bashevis Singer (1978). One Guild member, Elie Wiesel (1986), has won the Nobel Peace Prize.

immeasurable.

The book industry, happily, has to date fared better than our colleagues in the print media. This is, no doubt, partly due to our medium: print books are still superior in almost every way to their electronic counterparts. That advantage is rapidly fading, however; our transition to digital form is underway, and things change quickly in a digital environment.

The portents are not encouraging. Finding a sustainable business model for creative work in digital form seems nearly impossible on the Internet: if piracy doesn't get you, the aggregators will.

We'll likely need many things to go right to avoid the fate of our colleagues in the print industries.

The Opportunity

Yet, there are reasons for optimism. One of those reasons is our settlement with Google, which brings us here today. That settlement promises to address one of the oldest and most vexing of market failures: the loss to the commercial market of out-of-print books.

If you had asked knowledgeable people a couple of years ago whether we were close to delivering a near universal library to public libraries, colleges, and universities across the country, they would have scoffed. The technical challenges seemed too daunting, the rights clearance issues insurmountable, the passions stirred by the ongoing copyright wars far too intense for such a result to be achieved in the foreseeable future.

But here we are, on the cusp of that extraordinary achievement: the marriage of

much of our collective library with the Internet. The benefits to readers, students, and scholars would be profound. Here are a few:

1. The settlement would turn every library into a world-class research facility, by offering every public library building in the U.S. – all 16,500 of them – a free portal to millions of out-of-print books. The settlement would also offer a free portal to that same vast database of out-of-print books to more than 4,000 higher education institutions, from community colleges to our most elite universities.
2. Students and professors at colleges with the most modest of endowments would find an important part of the academic playing field had come to level, as they gain full access from every computer on campus to a library exceeding that of the finest Ivy League schools.
3. The visually and reading impaired would find the stacks of libraries open up to them as never before, gaining access to orders of magnitude more books than they currently have.
4. The settlement would offer anyone online in the U.S. free "preview" access to hundreds of millions of pages of text (up to 20% of each book).³ Readers from their own home computers would be able to review hundreds of accounts of the Battle of Vicksburg, or of the beginnings of the Industrial Revolution, or of the sources and interpretation of Moby Dick, at no charge.

³ Here's the math: we expect the settlement to make at least 10 million out-of-print books available, which, at an average of 300 pages per book, represents at least 3 billion pages of professionally written, professionally edited text. 20% of that is 600 million pages of text available at every desktop computer in the U.S. as a free preview. (For comparison, Encyclopedia Britannica is about 44,000 pages in print form; Wikipedia's featured articles total about 5,000 pages. All English Wikipedia articles, including stubs, total perhaps 3 million pages.)

Should a reader find one book particularly compelling, she could buy access to the entire book. Access to public domain books is free, of course, and authors controlling the rights to their books can choose to give away access for free.

Authors and publishers are willing to make this deal for several reasons. We of course hope to profit from the market that's created. We would like to have the Internet work for us, creating a market of the previously unmarketable. We also have a vital interest in keeping books central to our students, scholars, and culture. We're confident that making this vast library available online will help do just that.

Authors have another strong interest in making this deal work: authors need libraries. Libraries fuel their work. Authors of every type read, reinterpret and rely on their fellow authors, and those who have come before them. This is true of the scholarly writer and of the author of popular nonfiction. It's as true for authors of books for children as it is for authors of books for adults. Authors of literary fiction also rely heavily on those who've come before them. The creative expressions are new, but many of the ideas underlying literary works are eternal. Writers of genre fiction are no exception. Romance writers read romance novels and other works and offer their own interpretation and variations on the romance theme.

Authors, in short, want not only to realize the untapped value of their out-of-print works, they want access to this new, vast online library so that they can more easily create new works that readers will value.

The Opposition

Opposition to the settlement falls into several broad categories. We'll address two important objections in this Executive Summary.

Objection #1: Copyright doesn't permit a system that asks authors and publishers to specifically exclude their out-of-print works from uses negotiated on their behalf.

This simply isn't so. There is ample precedent around the world for dealing with market failures in copyright in precisely this way. For example, Germany today operates a system that nearly parallels the one the settlement would put in place. The German system allows for

- (a) routine copying of out-of-print works written by foreign and domestic authors,
- (b) including routine copying of "orphan" (unclaimed) works, and
- (c) without regard to whether those authors and publishers have expressly approved those uses.

The German system, however, actually denies authors and publishers any ability to exclude their works.

The German system goes even further, allowing the copying of in-print works without the permission of the author or publisher (in our settlement, the author and publisher must both approve of any displays of in-print works). Germany's photocopy licensing system is perfectly legal, and meets with the norms of international copyright law.

In fact, this is the typical way for countries to deal with the market failure represented by the unlicensed photocopying of copyrighted materials, although many countries allow authors and publishers to exclude their works from such licensing. Other countries with similar photocopy licensing systems include Australia, Canada, the United Kingdom and the Nordic countries, among many others. The market for photocopy licensing often fails without intervention, because the transaction costs of the license,

including the labor costs of the licensee, are simply too high relative to the value of the individual copy.

The inability to license out of print works to colleges, libraries and individual users presents a market failure on an epic scale. Here, as with photocopy licensing, a major component of the transaction costs involve rights clearance issues. Another impediment is the sheer scope of the project, and the capital and technological resources it demands. This settlement, with a financially strong and sophisticated technology partner, addresses the market failure. The societal value in bringing these works back to the market is incalculable, but until our settlement there was no practical way to do so.⁴

⁴ The impediments to making systematic use of the digital rights in out-of-print books are many, but they fall into two broad categories:

- A. Rights clearance issues. Here, there are three different obstacles. First, it may be unclear who controls the rights because no one knows whether rights to a particular out-of-print book have reverted to the author or not (most standard trade book contracts ask an author to demand a reversion of rights before the contract formally terminates and the rights revert to the author). Second, it may be unclear who controls the digital rights for a particular out-of-print book, since many (but not all) older contracts make no mention of digital or analogous rights. Third, it may be difficult to find the author or publisher who controls the rights, particularly for older works. This is the so-called orphan works problem. (More on this later.)
- B. Digitization and presentation issues. Here, the obstacles are capital and technological sophistication.

The rights clearance issues are largely addressed in Attachment A to the settlement, which deals with author-publisher issues and was the result of laborious negotiations between author and publisher representatives, and through the operation of class-action law.

The digitization and presentation issues have been handled with money and clever technology. Estimates of Google's costs in scanning, digitizing and building the technical infrastructure to support the display of millions of out-of-print books run to about a billion dollars. (We have no special knowledge of Google's costs.) Part of the challenge for Google was to find an efficient way to scan library books without damaging them by flattening them onto the scanner's surface. It solved this through a patented technology that corrects for the distortions caused by the scanning of a book page that curves away from the scanner's surface towards the book's binding.

Objection #2. The settlement inappropriately permits the use of unclaimed (orphan) works.

No issue has been more misunderstood or misrepresented regarding this settlement than the unclaimed or “orphan” works issue.

The primary misconception is the size of the problem: it’s much smaller than has commonly been reported, for several reasons. First, finding the rights owner of a book is not as daunting as many seem to believe. Books do not present the classic orphan works problem, photographs do. Photographs, both in the physical world and online, often become separated from their identifying information. This makes finding the rights owner a near impossibility. Books, however, always contain author and publisher information, and there’s often a copyright registration record to help locate the rights owner. Second, although a copyright-protected book may have been published as long ago as 1923, the vast majority copyright-protected books in our libraries are far more recent.⁵

Another major misconception is the failure to recognize that countries around the globe are already dealing with the orphan works issue in a productive way. The photocopy-licensing systems in other English-speaking countries permit the use of orphan works. This, as previously discussed, is a natural result of those nations’ attempts to cope with the market failure represented by unlicensed photocopying of copyrighted works. As the licensing societies collect photocopy royalties and start to cut checks to

⁵ This is because, for books published between 1923 and 1963, authors had to file renewal registrations to prevent their works from falling into the public domain. More importantly, the number of titles produced by the book publishing industry was far lower through most of the 20th Century than it was in its concluding decades. The median age of a copyright-protected book in a U.S. library is far younger than many commentators assume.

authors, word spreads, quickly, and authors step forward to register themselves. The pool of unlicensed works shrinks. The licensing societies are duty-bound to actively seek out authors⁶ for whom they have money. Year by year, they locate more and more authors, and the orphan works problem diminishes further.

We have some experience with this, since we helped found and long provided financial support to the Authors Registry, an independent, non-profit, rights-payment agency. The Authors Registry collects photocopy and other use fees from overseas, particularly from photocopy uses in the U.K., and pays authors in the U.S. the amounts due them. A sample of our success in paying authors of out-of-print works last year suggests that we reach 85% of such authors. The success rate of larger, more developed systems – such as that of the Authors' Licensing and Collecting Society (ALCS) in the U.K. – demonstrates that even higher success rates are possible. (The ALCS, representing more than 30,000 published writers, is an enthusiastic supporter of the settlement, for good reason. They know it can work, because they've achieved great success finding and paying photocopy revenues to authors of out of print books.)

Thus, this settlement presents a practical solution for the problem of orphan works for books. When an author is identified, then requests for all kinds of other uses – for permission to use an excerpt from the author's work, reprint it, or to translate it into a foreign language – can be relayed to the author or the author's agent and acted upon.

The orphan works issue is far smaller, and far more tractable, than some objectors would have you believe. The settlement itself is a big part of the solution.

⁶ Publishers are sought out too, of course. In some systems, such as in the U.K., different societies represent authors and publishers. In others, such as in Australia and Canada, the same society represents both authors and publishers.

Conclusion

We urge this committee to recognize this settlement for what it is: the outstanding result of a rare and productive truce in the copyright wars, negotiated by strong-willed and pragmatic representatives of the author, publisher, and library communities and a sophisticated technology partner.

To a dispassionate observer, we believe the solution presented by this settlement is how a rational, useful market for out-of-print books should operate in the digital age. The means of getting there, a class-action settlement, may be novel, but that shouldn't distract us from the great good – for readers, students, scholars, authors and publishers – that this settlement accomplishes. Similar systems, inevitably, will develop around the world.

This settlement doesn't pre-empt congressional action, but there's no need to act now, before we see how well this solution works in the real world. We suspect many of the concerns – including all of the major objections – will prove unwarranted as this settlement goes into operation. There's no need to fix that which likely isn't broken at all.

Allowing this opportunity to slip through our grasp would be a tragic loss to all those who value the riches stored in our nation's libraries.

II. INTRODUCTION TO THE SETTLEMENT

In 2004, Google Inc. (“Google”) announced that, as part of its Google Library Project (or “GLP”), it would reproduce millions of copyrighted books located in U.S. libraries and display “snippets” of those books on its website. These books had been published within and outside the United States and their copyright owners included

publishers and authors domiciled in both the United States and around the world. In copying and displaying content from books as part of the GLP, Google did not seek the permission of the copyright owners of those books.

The Settlement resolves two copyright infringement actions brought against Google by authors and publishers that alleged that Google's unauthorized copying of books and display of snippets without permission constituted copyright infringement under the U.S. copyright laws. Google defended its actions as non-infringing fair uses, as authorized by 17 U.S.C. § 107.

At the end of October 2008, after a negotiation that lasted more than two years, the author and publisher plaintiffs and Google announced that they had reached a comprehensive settlement of the litigation. The Settlement avoids further discovery and litigation, which could have been protracted.

The Settlement was carefully negotiated to protect the copyright interests of, and to provide meaningful benefits to a class of authors (and their heirs) and publishers of books covered by the Settlement (essentially, books published on or before January 5, 2009). One of the Settlement's crowning achievements for the class -- and for the United States public -- is that it breathes new life into millions of out-of-print books, which until now have been relegated to the dusty stacks of university libraries. The Settlement also provides new marketing and revenue opportunities for rightsholders of in-print books (should they choose to take advantage of them), in a way that does not harm the existing markets for those books.

If approved, the Settlement will provide the following material benefits to authors and publishers:

- **Payment to Rightsholders:** Google will be authorized to use class members' works in several revenue models, and will pay 63% of the revenues earned from its exploitation to a Book Rights Registry (the "Registry"), for disbursement (after an administrative charge) to the Rightsholders of those works;
- **Rightsholders Retain Control:** The Settlement is non-exclusive and does not involve any transfer of copyright ownership interests or any other property interests to Google. Additionally, Rightsholders at all times will retain control over their works, with the ability to determine the extent to which their works are to be included in or excluded from Google's uses, and to license their works to others;
- **Compensation to Rightsholders for Past Alleged Infringement:** Google will pay at least \$45 million to compensate class members whose works Google digitized without permission before May 6, 2009.
- **Establishment of Registry:** Google will pay \$34.5 million both for notice and settlement administration costs and to fund the Registry, a non-profit organization that will be managed by a Board representing authors and publishers, that will locate rightsholders, maintain a database of their contact information, collect and pay revenues to the class for Google's use of copyrighted works through this Settlement, and otherwise protect and represent the interests of the class.

As explained in more detail below, the Settlement provides extraordinary and previously unattainable benefits to authors and publishers in the United States and all over the world, along with remarkable benefits to the reading public, students, scholars and researchers in the United States.

III. THE MATERIAL TERMS OF THE SETTLEMENT

A. The Settlement Class and Sub-Classes.⁷

The Settlement Class is defined as follows:

All persons or entities that, as of January 5, 2009, have a Copyright Interest in one or more Books or Inserts. All Settlement Class members are either members of the Author Sub-Class or the Publisher Sub-Class, or both. Settlement Agreement ("SA") § 1.142.

⁷ Capitalized terms used have the same meaning as in the Settlement Agreement.

- **Author Sub-Class**

All members of the Settlement Class who are authors, and their heirs, successors, and assigns, and any other members of the Settlement Class who are not members of the Publisher Sub-Class. SA § 1.14.

- **Publisher Sub-Class**

All members of the Settlement Class that are (a) companies that publish books, and their exclusive licensees, successors, and assignees, and (b) companies that publish periodicals and have a Copyright Interest in one or more Inserts, and their exclusive licensees, successors, and assignees. SA § 1.120.⁸

B. The Parties.

The Representative Plaintiffs for the Author Sub-Class are authors Paul Dickson, Joseph Goulden, Daniel Hoffman, Betty Miles, and Herbert Mitgang. SA § 1.125. The Representative Plaintiffs for the Publisher Sub-Class are The McGraw-Hill Companies, Inc., Pearson Education, Inc., Penguin Group (USA) Inc., Simon & Schuster, Inc., and John Wiley & Sons, Inc. *Id.* The Authors Guild and the AAP are “Associational Plaintiffs.” SA § 1.12. The Associational Plaintiffs participated in the litigation to advance the rights and interests of the Author Sub-Class and Publisher Sub-Class, respectively, and served as consultants to counsel for the respective sub-classes.

⁸ A “Copyright Interest” refers to a United States copyright interest to the extent implicated by a use covered by the Settlement Agreement, i.e., reproduction of a Book and the Display Uses and Non-Display Uses (described below). SA § 1.38. A “Book” refers to a written or printed, bound work that was published or distributed prior to January 6, 2009 and, if a “United States work” under 17 U.S.C. § 101, registered with the United States Copyright Office. The Settlement excludes from the definition of Book works that are Periodicals, unbound personal papers, and works in the public domain (including government works). SA § 1.16. An “Insert” refers to copyrighted textual (but not pictorial) content in a book whose rightsholder is different from the book’s rightsholder. SA § 1.72. *See* the sections of the SA cited in this footnote for more details concerning these definitions.

Google was the sole defendant in the cases. In addition, though not parties to the litigation, the libraries of the University of Michigan, Stanford University, and the University of California took part in the settlement negotiations. They, along with a large number of other libraries, have entered into digitization agreements with Google, to allow their collections to be digitized. Numerous university libraries plan to participate in the GLP as well. A list of libraries that are authorized to provide books to Google for copying, and to receive digital copies, can be found at Attachment G of the SA.

C. Benefits To Authors and Publishers.

1. Revenues From Google's Use Of Books And Inserts

The Settlement authorizes Google, on a non-exclusive basis, to digitize Books and Inserts, to develop a searchable electronic books database, and to display and make commercial use of Books and Inserts, including the following "Display Uses":

- sell institutional subscriptions to schools, corporations, and government offices ("Institutional Subscriptions") (SA § 4.1);
- sell online access to Books to consumers ("Consumer Purchases") (SA § 4.2);
- display "previews" (up to 20% of a Book (SA § 4.3)) and "snippets" (several lines of a Book (SA § 1.147)) in order to spur Book sales and earn advertising revenues;
- place advertisements next to Book text (SA § 4.4); and
- provide free Public Access Services to all public and school libraries, with any printing from these services subject to a per-page fee (SA § 4.8).⁹

⁹ Also, subject to agreement with the Registry, Google may, in the future, be authorized to exploit works in other revenue models, such as consumer subscriptions, print on demand, custom publishing, .pdf downloads, and summaries, abstracts and compilations. SA § 4.7.

Note that all authorizations, including the right to digitize books and to display and make commercial uses of Books and Inserts, are confined to the United States and Google obtains no rights under the Settlement under any law other than United States law.

Revenues earned from Display Uses will be split 63/37 in favor of the Rightsholders. SA § 4.5. (“Rightsholders” are members of the Settlement Class that do not opt out of the Settlement.) Google will pay 63% of these revenues to the Registry, for disbursements to the Rightsholders.

2. Classification of Books.

The Settlement differentiates between Commercially Available (generally, in-print) and not Commercially Available (generally, out of print) Books. The purpose of this distinction is to ensure that Google’s Display Uses do not cannibalize the existing markets for Commercially Available books. Thus, the Settlement treats Commercially Available Books differently from Books that are not Commercially Available (i.e., Books that are not being marketed or sold by the publishers of the Books). SA §§ 3.2-3.3. Specifically, Google may not make any Display Uses of Commercially Available Books – both the revenue models described above and “snippet display” – unless both the author and publisher authorize Google to do so. *Id.* In contrast, Books that are not classified as Commercially Available will, by default, be included automatically in all Display Uses. At any time, however, the Rightsholder of a Book can change the default and instruct Google to exclude any Book from any one or more Display Uses. *Id.*

Google will initially determine whether or not a Book is Commercially Available as of January 5, 2009 (the Notice Commencement Date). SA § 3.2(d) (referred to in the SA as the “Commercially Available” classification). Google and the Registry will

continue to be assessing whether Books are Commercially Available using multiple sources of information. Rightsholders who claim their Books will be informed if a change is made in the Commercially Available classification of those Books.

Rightsholders can challenge the determination of whether a Book is Commercially Available, and any disputes with Google will be resolved through the Settlement's arbitration procedures (*see* SA Article IX). SA § 3.2(d)(iv).

3. Rightsholders' Do-Not-Digitize/Removal Rights.

Rightsholders can tell Google not to digitize their Books, or if already digitized, to remove their Books from Google's database and from the digital copies provided by Google to participating libraries. SA §§ 1.124, 3.5(a)(i). Google and the libraries must honor all do-not-digitize/removal requests made by April 5, 2011; thereafter requests will be honored only if the Books have not already been digitized. SA § 3.5(a)(iii).¹⁰

4. Rightsholders' Exclusion Rights.

Even if they do not remove a Book from Google's and the libraries' databases, Rightsholders at any time can exclude their Books (or portions thereof) from any or all Display Uses. SA § 3.5(b). Rightsholders at any time can exclude Inserts (or portions thereof) from all (but not less than all) Display Uses. *Id.* Rightsholders can change their mind at any time and turn Display Uses back on for some or all of their Books and Inserts. *Id.*

5. Rightsholders' Pricing Rights for Consumer Purchase

Rightsholders have two options under the Settlement for setting the sale price of their Books made available in the Consumer Purchase revenue model: they can set the

¹⁰ Exercise of the removal right does not vitiate Rightsholders' eligibility for the Cash Payment, described below.

price themselves (Specified Pricing) (SA § 4.2(b)(i)(1)) or they can rely on the “Settlement-Controlled Price” developed by Google. (SA § 4.2(b)(i)(2)). Settlement-Controlled Prices will be algorithmically designed by Google to find the optimal revenue-maximizing price for each Book. (SA §§ 4.2(b)(i)(2) & (c)).

Plaintiffs have heard from a number of academic and other authors (e.g., university professors) who are interested in making their books freely available to others through the Settlement. The Settlement Agreement enables Rightsholders to do so. The Settlement Agreement will support Creative Commons licenses. In addition, Rightsholders can set the sale price for their Books at zero. In those situations neither Google nor the Registry will receive revenues from any Consumer Purchase of those Books.

6. Establishment of The Book Rights Registry, and Payment for Notice and Claims Administration.

Google has agreed to pay \$34.5 million to fund the newly formed Book Rights Registry, to pay for notice to the Settlement Class and to pay for claims administration services. SA § 2.1(c). The Registry will be an independent not-for-profit entity, and all funds received by the Registry will be for the benefit of Rightsholders. SA § 6.2(a). The Registry will have a Board of Directors composed at all times of an equal number of Author Sub-Class and Publisher Sub-Class representatives. SA § 6.2(b). Thus, the Registry will not, as some have suggested, be controlled by Google. The Registry will establish and maintain a database of Rightsholders’ contact information and information regarding Rightsholders’ Books and Inserts. SA § 6.1(b). The Registry will locate Rightsholders, identify and coordinate payments to Rightsholders, and otherwise represent the interests of Rightsholders under the Settlement Agreement. *Id.* The

Registry will be able, with express rightsholder approval, to do deals with anyone, including competitors of Google.

The Registry will be funded at first by Google's \$34.5 million payment (net of notice and claims administration expenses). SA §§ 1.6, 2.1(c), 5.2. Thereafter, the Registry will also be funded by an administrative fee taken as a percentage of Rightsholders' revenues. SA, Attachment C (Plan of Allocation) § 4.2.

7. Payment For Already Digitized Books.

To settle plaintiffs' claims for actual infringement, Google has agreed to pay Rightsholders a minimum of \$45 million for Books and Inserts digitized without permission as of May 5, 2009 (i.e., the original opt-out deadline). SA §§ 2.1(b), 5.1. For every such work that is claimed by January 5, 2010, Google will make a "Cash Payment" of at least \$60 per Principal Work, \$15 per Entire Insert, and \$5 per Partial Insert.¹¹

If more than \$45 million is required to pay all of the eligible claims for Cash Payments, then Google will pay the additional funds necessary to make all such Cash Payments, with no cap on such additional payments. SA § 5.1(b). If the total amount of all eligible claims is less than \$45 million, the claiming Rightsholders can receive up to \$300 per Principal Work, \$75 per Entire Insert, and \$25 per Partial Insert. Any remaining funds thereafter from the \$45 million will go to the Registry for operations. SA, Attachment C (Plan of Allocation) § 3.2.

¹¹ "Principal Work" refers to a book's principal written work (such as a novel, or a collection of short stories). SA §1.111. "Entire Inserts" are complete works, such as short stories, forewords and poems. "Partial Inserts" are all other content that meet the definition of an Insert, such as quotes from other works. SA §§ 1.50, 1.100.

D. Author-Publisher Procedures.

Because, for most Books, authors and their publishers each have copyright interests, the authors' and publishers' respective interests are addressed in the Author-Publisher Procedures (SA, Attachment A). Among other things, the Author-Publisher Procedures set forth the authors' and the publishers' respective rights concerning in-print and out-of-print Books under the Settlement.

1. In-Print Books.

With respect to in-print books (other than works-for-hire), both the author and the publisher must agree to include the Book in the GLP; if they do not agree, it will not be included. Author-Publisher Procedures ("A-P") § 5.1. Cash Payments for in-print Books, as well as all revenues earned from Google's future uses of in-print Books, will be paid to the publishers of the Books, who will then pay the authors under the book publishing contract between the author and the publisher. A-P § 5.5. If an author wishes to challenge the revenue split offered by the publisher under that contract, the dispute may (except for Educational Books) be resolved in the Settlement's arbitration process. *Id.* The A-P also sets forth the respective rights of authors and publishers as to exclusion, removal, pricing, and changes in Display Uses of in-print Books. A-P §§ 5.2-5.4.

2. Out-Of-Print Books.

The A-P includes provisions concerning exclusion, removal, control of pricing, and changes in Display Uses of out-of-print Books, which vary depending on whether the Book is (1) a work-for-hire under United States copyright law (100% of Cash Payments and revenues paid by the Registry directly to the publisher), or (2) reverted (100% of Cash Payments and revenues paid by the Registry directly to the author). A-P § 6.1.