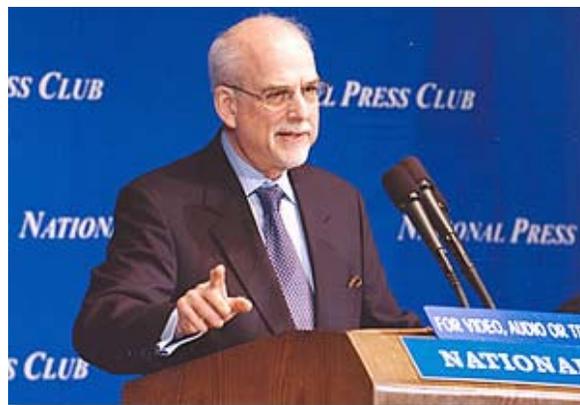


## Speech by Chairman Bruce Cole

## NATIONAL ENDOWMENT FOR THE HUMANITIES

National Press Club  
November 16, 2004  
"30 Million Pages to Go:  
Digitizing the American  
Newspaper"  
(As Prepared for Delivery)



Thank you so much, Sheila.

I've got a confession to make about newspapers. Aside from reading papers, my only real-life connection with them was as a very small business man.

More decades ago than I care to recall, I delivered the *Cleveland Plain Dealer*. I was proud of my new white canvas bag emblazoned with the name *Cleveland Plain Dealer*. And I was equally proud of the neat money changer fixed to my belt.

This was a tough job, trudging up steep driveways in the arctic winter of Cleveland, the fear of dogs, imagined and real, always ready to take a bite out of a carrier. And the customers were a little fearsome too, especially if they didn't get their papers on time. They valued their newspapers and could be extremely put out if I was just a tad late. This early experience gave me an understanding of how important newspapers are.

One of the distinguished publishers in this town, Philip Graham, called newspapers "the first rough draft of history." That's an apt definition--and that's what we're working with. For more than 20 years, we have been saving that history through the United States Newspaper Program--preserving the newspapers that date back to our earliest days.

These papers have wonderful names--*The Cain County Razooper* from Kansas, *The Daily Unterrified Democrat* from Colorado, *The Georgia Temperance Crusader*, and from my own home state, *The Castigator* of Ripley, Ohio. We have rescued the information in a quarter of a million American newspapers like these, newspapers that are crumbling because of acid paper, and poor storage, or simple neglect.

If you've ever left a newspaper out on a doorstep for a few days and come back to find it yellowing at the edges, you know what I'm talking about.

We are now launching a new effort--the National Digital Newspaper Program with our partners of the past 20 years, the Library of Congress. The Librarian of Congress Jim Billington cannot be with us, but Deana Marcum, the Associate Librarian, is here today, along with Mark Sweeney, the guru of the Library's newspaper program. And so is the man who has been Mr. U.S. Newspaper Program at the Endowment, George Farr.

Now we are embarking on the exciting next step. We have already microfilmed 67 million pages of newspapers. With the Library of Congress, we will begin to digitize 30 million pages. Anyone who's interested--students, historians, lawyers, politicians--

even newspaper reporters--will be able to go to their computer at home or at work and at the click of a mouse get immediate, unfiltered access to the greatest source of our history. You will be able to search in day-to-day accounts in these old newspapers. The project will be based at the Library of Congress--and the material will be available to the American public for free, forever.

So what does this treasure consist of? Some are events we know about. The Chicago Fire. The discovery of gold. The San Francisco earthquake. The Battle of Gettysburg. But the papers tell us more. They give us what lies beneath the headlines--the ordinary daily record of life. They provide us whimsical moments, or tragic ones or hilarious ones--they are all there on the pages. Let's take one day in history.

Let's pick a day. Today, November 16, nearly 100 years ago, 1907. Oklahoma became a state. The headline on the *Tulsa Daily World* that day says "Roosevelt to shove the quill at nine," meaning he would sign with an eagle quill pen. There are other little bits. The governor of the Oklahoma territory refused to ride in the inaugural parade. He was mad. "I am not about to talk for publication. I simply do not want to take any part whatever. . ."

There's more. Here's an ad from the *Pure-Food Grocer*. A poem, to be more precise.

"For Tulsa is a growing town  
And demands stores that's good.  
We couldn't run a dirty place,  
And wouldn't if we could."

Or Dr. Pierce's Favorite Prescription. It would be an advertorial today. Dr. Pierce's, says the article, is "the only medicine designed for the cure of . . . peculiar ailments. . ." and, quote "contains no alcohol, and no narcotics, or other harmful or habit-forming drugs."

This is not the condensed stuff of standard history textbooks. These old papers give us an eyewitness view of history. You can go to their pages and read about weddings and births, about McGuffey spellers for sale, about gossip, about what people were eating and drinking, about almost everything. Newspapers give us glimpses of the economic, political, commercial, and social dimensions of our country. They give us a view that at the same time offers the details and the panorama.

Here's why newspapers are so valuable. A democracy like ours is only as good as the knowledge of the people who are part of it. The more we remember of our past, the better off we are--the stronger we are. But tests and surveys have shown that our citizens, especially our young people, exhibit an alarming lack of knowledge about their history.

The historian David McCullough has expressed concern that the present generation needs more awareness of its past. He said: "I think we are raising a generation of young Americans who are, to a very large degree, historically illiterate." He says that in times of crisis "we should draw on our story; we should draw on our history as we've never drawn before."

Knowledge of our history is not a luxury, it's a necessity. American Amnesia is dangerous. Democracy is not self-sustaining; it needs to be learned and passed down from generation to generation. We have to know our great founding principals, how our institutions came into being, how they work, what our rights and responsibilities are. As the founding legislation of the NEH says, "Democracy Demands Wisdom."

Our Founders were well aware of this. Just after Benjamin Franklin signed the Constitution he was asked what the Founding Fathers had given the people--a monarchy or a republic. Franklin answered, "A republic--if you can keep it." We've

got to keep it by knowing our story, the whole story, the center and the margins, the good and the bad, for it's a remarkable story.

As President Bush has said: "Our history is not a story of perfection. It's a story of imperfect people working toward great ideals."

At the NEH we've embarked on an initiative called *We the People*, which addresses our historical amnesia. It's supported by the President and Congress--and last year we were given a record budget increase to launch it.

Here's how it works. By developing and strengthening programs on American history and culture, NEH is providing the nation with resources for exploring and understanding our collective past. We're helping teachers expand their understanding of U.S. history so they can return to the classroom with greater depth. We're collecting the papers of U.S. presidents. We're supporting museum exhibitions on American culture. We're providing forums for citizens to learn about U.S. history--from exhibitions to reading and discussion groups. And we're making books for kids available to libraries across the country.

The new National Digital Newspaper Program is a cornerstone of this effort. Newspapers are a singular source for understanding the fabric of the towns and regions of our country. 1846. *The Californian*: the first newspaper in the state. Printed in both English and Spanish. Or *The Cherokee Advocate*: 1844. The first newspaper in the Oklahoma territory; two pages in English and two in Cherokee, carrying missionary and farming news. Or Chicago in 1900. Newspapers printed in German, Greek, Polish, French, Bohemian, Italian, Yiddish, Slovenian, Hebrew, Lithuanian, Danish, Norwegian, and Swedish. Plus English.

The stories in America's newspapers, no matter what language, offer little pieces of history, which combine to form the mosaic of our past, the stories that were news before they became history.

Now, with this new digital program, you will see the papers just as they were--you will be able to search the actual page. The technique is OCR--optical character recognition. In fact, there is already a model up on the Library of Congress site. It's got the *Stars and Stripes* from World War One. It shows you the whole page and there's a zoom device so you can focus in on a single story and be able to read it. It's key word searchable. It's a quantum leap from trying to read microfilm.

There are difficulties in putting any of these newspapers on the web. Newspapers have stories of different sizes; different typefaces; advertising; jumps--a story on page one that continues on page 15. It's complicated. It's going to take a couple of years to get the newspaper project up and running.

When it is complete, it will ultimately cover 1836 through 1922. That sounds strange, but it's for a couple of very good reasons. The type from colonial times has elaborate fonts and OCR technology can't read it well yet.

At the other end--by 1923--we run into modern copyrights.

Eventually, whether it's online or offline, the holdings and their locations will be part of a bibliography the Library of Congress is assembling. It will tell where every newspaper is located--from the first American newspaper in 1690 to the present day.

That first newspaper, by the way, was *Publick Occurrences* and it was published for exactly one issue. The governor of the Massachusetts colony shut it down for rabble-rousing--talk about freedom of the press! There is no copy of it in this country. As far as we know, the only one is in the Public Records Office in London--it was sent there as evidence of our wayward ways.

Meanwhile, we have more than enough to do as we go into digitization. In this new phase, every state will eventually be represented. We are starting with a test run of a million pages covering 1900 to 1910. It's an interesting decade--the San Francisco earthquake, the presidency of Teddy Roosevelt, the Wright Brothers, and, as we know, Oklahoma.

It was a high point in numbers of American newspapers. There were 2600 dailies in 1909 and 14,000 weeklies. And there's another reason that period is fascinating. It's when newspapers began to develop their modern look--photographs, maps, decked headlines, and so on. It will be a useful test, we think.

After we complete the test phase, we will be joined by a third partner, representing the states--state libraries, for instance, or state historical societies, or universities. We will invite them to lay out their plans and choose newspapers to be included in the digitization. We have three broad criteria.

1. High research value

Is it the paper of record?

Is it the one that carries legal notices?

2. Geographic spread (is it multicounty)

And,

3. Chronological span--Was there something going on that was important--an event, a movement?

This new technology is transformational. I remember the days when I had to go to a research library, or wait for interlibrary loan, or spend hours reeling through microfilm. The microfilm will still be there as backup, if needed, but the search will be from home and the answers as near as your computer.

This digitizing will democratize knowledge by making it available to anyone with an internet connection. But just as important and revolutionary, it is also going to create something new. The sheer volume of information in newspapers has been an obstacle. Newspapers carry 3000 to 7000 words on a page. The new technology overcomes that. The page is scanned; it's tagged with name, date and page number--metadata. The process turns the enormous volume of material into a searchable asset. And this asset will be easy to use.

By being able to search and sort the metadata, we will be able to ask new and more sophisticated questions, which will create new knowledge. It's still out on the horizon but we can see it coming.

What will happen in the future? What are the successors to newspapers? Clearly, magazines and newspapers have co-existed for a long time now. And television hasn't driven out newspapers.

But what about the new democracy of the Internet? Now we have the Web and the millions of eyes and the voices of the bloggers. Are they our Tom Paines and the pamphleteers of our digital revolution? What's going to happen to material like theirs which is created electronically? Should we be thinking about preserving that? And what about manuscripts, emails, and so much else created without paper?

As I'm sure you've noticed, I'm providing more questions than answers at the moment. I'm a historian, an art historian to be precise, not a newspaper editor, and I can't claim to be an expert in computer technology. But I do know that the National Digital Newspaper Program, like the United States Newspaper Program before it, is

one of the most monumental and significant projects undertaken by the NEH for the benefit of the American people. It is a platform of fact, concept, and knowledge. It will inform, entertain, enlighten, and instruct our citizens for as long as we can imagine. We are proud to be playing a part in it.

Thank you very much.