



**PUBLIC.RESOURCE.ORG** ~ *A Nonprofit Corporation*

---

**Open Source America's Operating System**

"It's Not Just A Good Idea—It's The Law!"

August 12, 2009

Congressman Michael Honda  
1713 Longworth HOB  
Washington, DC 20515

Dear Congressman Honda:

I was writing to let you know that the Government Printing Office has notified me that the Official Journals of Government will be available at no charge in bulk as XML source code, starting with the Federal Register in September. I believe your leadership in this area, particularly the Honda Bulk Data Amendment to the Legislative Appropriation Bill played a large role in convincing the Government Printing Office and the Office of the Federal Register that the Congress would support aggressive steps by the agencies to open source these key databases.

In the original requests to make bulk data from the executive branch available to the public, in addition to the Official Journals of Government, we had asked for changes in access to U.S. Patents, a database specifically mentioned in the U.S. Constitution as a resource to "promote the progress of science and useful arts."

Public.Resource.Org maintains a collection of the full text of patents from 1983 to 2000, but if we want to "top off" our collection and make it current, that would cost \$6,000 per year and a whopping \$39,000 for the current-year subscription. Full access to the patent database, in the form of binders that include the application, grant, appeals, and other materials surrounding a specific patent, would cost orders of magnitude more.

I've been trying since 1994 to get U.S. patents available on the Internet in bulk so anybody can download this constitutionally-mandated database and thus promote progress. Yet, to provide such a resource would cost us \$100,000 or even more. This does not seem to be in keeping with the spirit of the Constitution, nor of more prosaic regulations such as OMB's Circular A-130 on distribution of government databases.

Any assistance you might provide in bringing this issue to the attention of Commissioner Kappos or others in a position to consider a change in information distribution policies would be greatly appreciated. Again, thanks for your leadership in helping make the Official Journals more broadly available.

Sincerely yours,

Carl Malamud  
President & CEO  
Public.Resource.Org

June 25, 1998

## U.S. to Release Patent Data On a World Wide Web Site

By JOHN MARKOFF

In a new plan for making Government information freely available over the Internet, the Clinton Administration will announce today that it will make the full data base of the nation's patents since 1976 and trademark text and images starting from the late 1800's available on the World Wide Web beginning in August.

The project, which will create the largest Government data base on the Internet, is to be announced by Bruce A. Lehman, Assistant Secretary of Commerce and Commissioner of Patents and Trademarks, in a speech to the American Bar Association's section on intellectual property law in Williamsburg, Va.

The decision to make the data base freely available is a result of a fierce debate that has gone on for years between public interest advocates who argue that Government information should routinely be made available on the Internet and companies that purchase the data from Government agencies to resell.

Indeed, the Clinton Administration's action came less than a week before a deadline imposed by Carl Malamud, an independent Internet pioneer, who in May said he planned to purchase the data from the Patent and Trademark Office and make it publicly available if the Government failed to act.

After imposing the deadline in a letter to Vice President Al Gore Jr. and the Commerce Secretary, William M. Daley, Mr. Malamud said that he had received an anonymous donation to support his project and that he was developing his own system of patent and trademark records.

Mr. Lehman, however, said that his agency had already made a commitment to Mr. Gore to make the data base available as part of the Vice President's "reinventing government" program.

Mr. Malamud said today that he had decided to shelve his plans.

"Our site was a backup in case the Administration refused to budge," he said. "We very pleased with the outcome of this five-year struggle. The American public gets the data they deserve, and the Clinton Administration has shown real responsiveness by reversing its earlier policy."

In January 1994, Mr. Malamud's group, the nonprofit Internet Multicasting Service, posted the full text of filings made by corporations to the Securities and Exchange Commission and several years of the patent data base over the objections of the S.E.C. and the Patent and Trademark Office. His organization later added data bases from the General Services Administration, the Federal Election Commission, the Federal Reserve Board and the Government Printing Office.

In a telephone interview, Mr. Lehman said that it had long been his intention to make both data bases available on the Internet but that he needed first to meet a range of technical and policy criteria.

"Creating a 1.3-terabyte searchable data base is a big technical feat," he said. A terabyte is one trillion characters of information, equivalent to about a million copies of "Moby Dick."

First, he said, such a system requires tight security provisions to isolate it from the Patent and Trademark

Office's internal computer network. In addition, he said, his agency had to perform an economic analysis on how the decision would affect both the Patent and Trademark Office and the information industry.

"We have no interest in competing with the private sector," he said, adding that his agency's analysis had shown that the private-sector offers added value beyond the patent and trademark documents.

The new data bases will be available at the Patent and Trademark Office's Web site beginning in August, when trademark text will go on line. Trademark images and patent text will follow in November. Patent images linked to the text of the patents will be made available by March 1999, he said.

The data base of more than two million patents will be searchable by key word, as will more than 800,000 registered trademarks and 300,000 pending trademarks. The entire data base will be made up of 21 million documents.

Currently the Patent and Trademark Office makes abstracts of patents available, a system that is already providing users with more than three million pages of patent material a month.

The Patent and Trademark Office received 237,045 patent applications in the most recent fiscal year -- a 14.9 percent annual increase -- and approved about 114,000 patents. Trademark applications grew by 11.8 percent, to 224,355.

The new data base on the World Wide Web is a significant step forward, Mr. Lehman said, adding that his plan calls for completely automating the nation's patent system by 2003.

# Ideas whose time for free access has come

By John Schwartz | *The Washington Post*

Friday, July 03, 1998

---

If the World Wide Web could be felt and not just merely surfed, you'd be feeling a very big rumble in August. That's when technicians will begin connecting one of the largest single databases ever offered on the Web. It's the official record of the U.S. Patent and Trademark Office (PTO).

How big a deal is this? Bruce A. Lehman, commissioner of patents and trademarks, put it to me this way: "This database is the record of technology at this moment in time."

If some unimaginable holocaust were to zap the United States, he said, survivors could pull the PTO's backup tapes out of the Pennsylvania salt mine where they're stored and "entirely reproduce all of the technology of the 20th century." And now, he continued, "we are putting the entire library of the technology of our time on the Web, available with a few keystrokes."

If you're not impressed yet, you just might need to have another cup of coffee.

But before we say more about this new stuff, let's build up to it by discussing some very old stuff.

You might think this is a nation built of laws, or of power, or of money. But it's also a nation of ideas. This whole country, as Abraham Lincoln said at Gettysburg, was invented because of a "proposition": the once-crazy notion "that all men are created equal." To ensure the survival of the marketplace of ideas that helped them create the structure of the new government, the Founding Fathers made freedom of speech the Constitution's very first amendment.

Just as important, they enshrined within that Constitution the notion of protection for ideas that might make people money through a system of patents and copyrights. There it is, right there in Article 1, Section 8: Congress shall have the power "to promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."

Jefferson, Franklin and many other prominent Americans present at the birth of the republic were scientists and inventors on the side. They knew the power -- and profit -- that can come from an idea. But like all of the great notions at the core of our nation, the system of patents and copyrights they envisioned was a delicate balancing act: On the one hand, the person who comes up with ideas deserves a degree of protection. On the other, those ideas become infinitely more powerful when shared and built upon.

All right. Now let's click forward a couple of hundred years and many millions of patents later. I was sitting at my desk when the phone bleated its electronic tone. I picked it up and heard the unmistakably impish voice of Carl Malamud, one of the Internet's more provocative guys.

Malamud believes that government information belongs to all the people and ought to be freely available on the Web. He's on one side of a long-standing tug of war with companies that profit by packaging government information and reselling it to businesses and consumers.

Through his Internet Multicasting Service, he's had a hand in pushing an incredible amount of such information into the ether, most notably filings with the Securities and Exchange Commission, the Federal Election Commission and Government Printing Office. He's often helped force those government agencies to take over and improve the online operations. So whenever Carl calls, I figure I'm in for a pretty interesting ride.

This time he had his sights set on patents, one of the biggest kahunas out there. They've long been available by mail and at one of 70 national patent libraries, but he was agitating to get them online. He was calling me to say he was fed up with waiting. After Malamud made his frustration known through *The New York Times*, a Washington-area Internet entrepreneur anonymously ponied up \$500,000 and told Malamud that he should just do what big business does and buy the patent information himself -- and then give it away.

Malamud started working round the clock and began quietly telling journalists that his bare-bones patent Web site would open its virtual doors on July Fourth.

Then last week, Commissioner Lehman made his announcement. First will come trademark text libraries, in August, followed by trademark images and patent text in November. Patent images will begin appearing in March 1999. The full text of the 2 million patents dating back to 1976, along with trademarks from the 1800s onward, will be online, Lehman said, joining 20 years of patent abstracts and full AIDS-related patent databases that his office already makes available online.

That will be the start, the patent office says. The grand aspiration: that one day every single patent going all the way back to the beginning will be online.

The new databases will be searchable by key word -- a crucial feature that Malamud had no plans to provide, since he assumed others would jump in to fashion innovative tools for exploring the trove. "For the first time in history," Lehman told me, "patent info will really satisfy the intent of the Constitution." Anyone with access to the Web will be able to share the wealth of the world's creative genius; "Now just-ordinary people will be able to have that information at their fingertips."

Those who currently profit by reselling patent data will still be able to do so, Lehman predicts, by finding ways to help users draw needles of usable information out of that vast haystack of data.

Malamud is now withdrawing his Web patent effort, and he praises the Clinton administration for making this treasure available to us all. "It's a good thing," he said when we talked. "The intellectual-property market is going to change because of this."

Lehman denies that Malamud's campaign figured in what the PTO will now do. Putting the database on the Web site, he notes, has been in the works for some time and is a key element of both Vice President Al Gore's "Reinventing Government" initiative and Commerce Secretary William M. Daley's efforts to boost electronic commerce.

But before launching the site, the PTO had to first ensure that its patent reviewers had access to the full database, Lehman said, and then a separate system for the general public had to be developed to ensure security. These things always take longer than activists would like. "Like everyone, he wanted it yesterday," Lehman said of Malamud with a sigh.

However it happened, I'm just overjoyed that this vast data tsunami is going to hit the Net. I think about people like my pal Scott Campbell, a New York inventor who drives down to Washington a few times a year to do patent research, who now will be able to do much of that work from his home.

I think of high-school students who might someday be inspired by examining online the images of Thomas Edison's patent application for the first phonograph. I'm wondering what ideas will be sparked by their searches.

And I know that somewhere, the Founders are smiling.

Schwartz can be reached at [schwarti@twp.com](mailto:schwarti@twp.com)