CONTROL #: C540726
WHITE HOUSE CONTROL #: 
ACTION AGENCY: PTO
COORDINATE WITH: 
ADDRESSED TO: SECY
SIGNATURE LEVEL:
INSTRUCTION: FOR INFORMATION

FROM: MALAMUD, CARL
(IMS)
TITLE:
SALUTATION:
NUMBER OF COSIGNERS: 0
SUBJECT: REQUEST THE PUBLIC IS ABLE TO MAINTAIN ACCESS TO KEY DATABASES

INFO COPIES: ES,

RECEIVED: 08/08/95 ASSIGNED: 08/08/95

THE CORRESPONDENCE ANALYST FOR THIS DOCUMENT IS BD
BD
August 4, 1995

Hon. Thomas Daschle
U.S. Senate
Washington, D.C. 20515

Dear Senator Daschle:

You will find enclosed letters that were sent today to Chairman Levitt of the Securities and Exchange Commission and Commissioner Lehman of the Patent and Trademark Office. For the past 19 months, the non-profit Internet Multicasting Service has posted SEC and Patent electronic documents on the Internet, providing over 4.7 million documents free of charge to the American public. Over 20,000 documents per day are distributed through this "information safety valve," the only readily available public source for these key government databases.

On October 1, we will terminate this service. As a public service, our small non-profit, with corporate contributions, a National Science Foundation grant, and a joint venture with New York University, undertook this demonstration project. We wanted to show that there is a cost-effective, secure, technically effective way to get large government databases distributed to the American public. The project has clearly succeeded.

Our users range from college students looking for jobs to corporate engineers to senior citizen investment clubs to workers trying to track their mutual funds. The two-year demonstration project ends October 1, and we are concerned that neither the SEC nor the Patent office have taken any steps to ensure that a public source of data remains available.

Under the leadership of both Congressional Republicans and the Clinton Administration, the Congress and the President recently passed the Paperwork Reduction Act. This law makes it an obligation of agencies to ensure that their information be made available in an equitable manner to all citizens with a diversity of public and private sources.

On October 1, there will be no diversity and no equitable access. Let me give you a concrete example. Microsoft's annual report for 1994 is available on our system at no charge to the user. The largest commercial source for this data, Lexis-Nexis, sells the same document for $622. We certainly defend their right to sell documents at any price: that's the American way. But, it is only common sense that there must be alternate sources available for such important public information.

The SEC and Patent documents are basic enabling documents for our information economy. The purpose of the public disclosure requirements in the SEC Acts of 1933 and...
1934 is to guide investment dollars to the right portions of our marketplace. The very purpose of our Patent system is to encourage the rapid growth of technology and science by documenting the state of our knowledge. These databases are not products or profit centers, they are the very fuel of our information economy.

This issue is clearly nonpartisan. The Clinton Administration has repeatedly praised our efforts, calling the Patent project "a big win for the American public" and the SEC project "an Administration priority." The Contract with America made equitable and timely distribution of government information a priority. President Clinton, on the bill's signing, remarked on the fact that there was not a single dissenting voice in Congress for this important piece of legislation!

This issue is a key test of our resolve to build a National Information Infrastructure. Will the remarkable advances in our public policy towards new technology in the past 2 years turn into something real, or will the efforts of a few government bureaucrats block change? Will the new law be turned into an empty symbol of what might have been? Will public information become truly public or will we continue to auction America's databases to the highest bidder?

I hope you will take steps to ensure that the public is able to maintain access to these key databases in an equitable and timely manner.

Sincerely,

Carl Malamud
Dear Commissioner Lehman:

Since January, 1994, the non-profit Internet Multicasting Service has maintained an on-line database of U.S. Patent documents, as well as the full text of SEC Edgar documents, for the American public to access at no charge. This service will terminate on October 1, 1995.

Our on-line service allows keyword searches of the full text of U.S. Patents using the WAIS, World Wide Web, Gopher, and Electronic Mail services. Bulk transfers are available via the File Transfer Protocol. A digital signature is added to each document to allow users to verify the contents. Background material and links to additional resources are also maintained on the system.

Since January, 1994, we've sent out over 1.5 million Patent documents to the public. Over 2900 documents per day are now being distributed, and there is a strong demand by users for the full database (the more fully developed SEC system, for example, sends out over 17,000 documents per day). Our current database of patent documents is 180,000 files and 7.1 Gbytes. Given the current average rate of data, we estimate that a year of your data is approximately 4.5 Gbytes.

Based on the access figures and the size of your database, it would be a trivial task to put the entire Patent database available on the Internet. The cost of computer equipment, routers, and disk drives is well under $150,000. Our staff maintains the Patent dissemination service for well under 0.5 FTE of a systems programmer and 0.5 FTE of a developer, for a total labor cost of under $100,000. Internet access is under $25,000/year. If you depreciate the equipment over 3 years, the total cost per year of providing Internet access to your entire database is $175,000.

Our Patent and EDGAR service has been funded partially by a grant from the National Science Foundation. Additional support was provided by our corporate sponsors, by a coalition of university research centers, and with a great deal of "sweat equity" and our own money added by our dedicated staff committed to this important public service.

Recently, the Congress passed and the President signed the Paperwork Reduction Act, which becomes law October 1. This bill was a key provision of the Contract with
America and has received the strong support of the Clinton administration. The law reads:

"(d) With respect to information dissemination, each agency shall--

"(1) ensure that the public has timely and equitable access to the agency's public information, including ensuring such access through--

"(A) encouraging a diversity of public and private sources for information based on government public information;

"(B) in cases in which the agency provides public information maintained in electronic format, providing timely and equitable access to the underlying data (in whole or in part); and

"(C) agency dissemination of public information in an efficient, effective, and economical manner;

In case there is any doubt, the committee reports go on to stress that information dissemination is an "integral part of the information life cycle," and emphasize that the purpose of the law is to "enunciate clearly the obligation of Federal agencies to ensure effective public access to government information."

Many of us have learned from the last two years of this demonstration project. It has been remarkable to see how large a portion of the American public has interest in these key documents. It has also been remarkable to see the advances in technology that allow dissemination to all members of the public a realistic and cost-effective goal. We hope you will embrace both the spirit and the letter of the laws recently passed by the Congress and signed by the President and that you will ensure that Patent data remains available on-line and does not get auctioned off to the highest bidder. We look forward to seeing your Internet service and will continue to be available to your staff as you make this important transition to the information age.

Sincerely,

Carl Malamud

cc: Vice President Gore
Hon. Ron Brown, Secretary of Commerce
Ms. Sally Katzen, Office of Management and Budget
Hon. Newt Gingrich, Speaker of the House
Hon. Robert Dole, Majority Leader of the Senate
Hon. Connie Mack, Republican Conference Secretary
Hon. Thomas Daschle, Minority Leader of the Senate
Hon. Richard Gephardt, Minority Leader of the House
eral health study and ultimately every memo of the President's, scrawled or E-mailed. They all join a virtual public library — arguably the most valuable storehouse of data on the planet. It is also one of the most archival: scattered across departments, poorly cataloged and expensive to access. And, with few exceptions, it is off line.

Meanwhile, across the Potomac from the patent office, in a tiny room in the District of Columbia, sits the Internet Multicasting Service, which is not much more than a fast telephone line attached to a donated work station and disk drives. The service has obtained the patent data for last year and the first months of this year and put it on line, free, fully indexed and searchable. Carl Malamud, the service's founder, says he could just as easily offer the entire historical data base — but the office won't give him the data.

"If you treat this as a product, but it's not a product — it's enabling information," he says. It is, or should be, fuel for the information economy.

The patent office already has a high-bandwidth Internet connection. That could easily enable any of the millions of home and business computers with access to the Internet to plug into its system and see what a user sees at that Arlington terminal, just as any computer can now plug into the New York Public Library's on-line catalogue or the data bases of thousands of other libraries. The public has already paid more than $400 million to create a patent data base available only to walk-in traffic.

So why not go on line? The commissioner's responses echo the reasoning of scores of other government agencies, federal and local, facing the same issues:

* It's not our job. "We're not a library," Lehman, the commissioner, says. "It's not the fundamental purpose of the patent office. Now if Congress wants to change that they can, and they can provide an appropriation to do that."

* We're doing it anyway, as fast as we can. "It has always been part of our plan to provide a plug into the patent office to outsiders," Lehman says. "We are not at this moment in time in a position to open up that plug for technological reasons."

* And, we must not compete with the private sector. "We're developing a big information industry in the United States. We already see about 30 companies that feed off the patent office, and we want to encourage that. Part of what we're trying to do is bootstrap new industries."

That last argument sounds attractive, until you realize that those companies are lobbying for the privilege of paying the Government more — in other words, they want to forestall competition. They belong to an industry that has used heavy, targeted, legal contributions to protect its stake in an economic model that is rapidly becoming obsolete: scarce data sold to specialists at high prices. West Publishing, with a near-monopoly on the Government's court data bases, is a costly example, as lawyers quickly discover. The Internet has created a different model: information of all kinds, a mass audience, low prices.

Lehman acknowledges that private-information services lobby him hard to keep prices up; he denies being influenced by their pleas. Nevertheless, the patent office, like many other Federal agencies, sells its data mostly on old-style mainframe computer tapes, at prices low enough to guarantee enormous profit for commercial services but just high enough to prevent widespread distribution.

A potentially far-reaching new statute, passed with little fanfare this spring, requires Government agencies to make electronic data available for no more than the "cost of dissemination." Twenty-five years of patent data — for which the patent office charges nearly $200,000 — would fit on tape cartridges costing no more than a few hundred dollars. How to explain that? The office's arithmetic counts not just the tapes and the few technicians hours that would be needed to fill them with data, but also fuzzier items: general staff time, updating the data bases and "customer service."

This is an Administration that has javelined hard for the on-line world since taking office. The White House has an elegant home page on the Internet, with digital renditions of official seals, speeches and their press releases — political information sitting in for the real thing. Yet where the truly valuable data bases are concerned, the Clinton Administration has produced no comprehensive plan for the future. It may not be necessary. The White House could take a powerful step forward merely by leaning on its bureaucracy: ordering them not just to comply with the new public-information law but to embrace it. That would mean taking the crucial phrase cost of dissemination literally: cheap disks or tapes; duplicating and shipping and nothing else.

"People are concerned about universal access — the wire running into your house will be the easy part," says James Love, director of the Taxpayers Assess Project, a Washington advocacy group. "Certainly the one thing people shouldn't have to worry about is government information, the thing they own as taxpayers. There'll be lots of other things they won't be able to afford. At least this should be available."
Under this system, a retail information provider, like Mead Data's own Nexis service, charges about $15 for each S.E.C. document, plus a connection charge of $19 an hour and a printing charge of about $1 a page. The only fees to use the S.E.C.'s data base under the science foundation's project would be for access to the Internet, for which pricing varies. Commercial access can be bought for as little as $2 an hour. But many college students now obtain Internet access as part of their tuition and many businesses pay a high-speed Internet connection that might cost the company hundreds or thousands of dollars each month but permits employees to share unlimited access to the network.

A legislative-affairs lawyer for the Information Industries Association, Ronald Pless, said the industry would not oppose the project as long as the Government did not intend to restrict commercial publishing of Federal data. "We have no problem with data bases being made available over the Internet," Mr. Pless said. "But there has to be an interest in insuring a diversity of sources. We don't want a Government monopoly on the ownership and dissemination of Government information."

Representative Edward J. Markey, a Massachusetts Democrat who is chairman of the House Telecommunications Subcommittee, has been pushing for greater public access to the S.E.C.'s Edgar Dissemination Service. He said today that he applauded the science foundation's move. "I've been trying to launch Edgar into cyberspace for the last year, and I think the S.E.C. finally gets it," Mr. Markey said, adding that he expected the project to stimulate the commercial market rather than hurt it. "This project will prove to be a friend of the private information providers," he said. "It gives people a taste of what on-line information is." The company involved in the project, the Internet Multicasting Service, is a nonprofit organization founded by Carl Malamud, an ecologist who developed computer technology for the Federal Reserve Bank. Mr. Malamud has also been instrumental in creating technology that is capable of broadcasting information including audio, video and data over the Internet. "We're not interested in replacing Mead Data," Mr. Malamud said. "I'm not in the financial data business. This is designed to democratize the distribution of financial information."

He said that the data would be delayed by a day, in contrast to the instant access that is provided by some on-line publishers for financial professionals. Under the terms of the science foundation grant, the New York University researchers will buy raw data and reform it so that it can be obtained easily over the Internet through Mr. Malamud's organization. Mr. Malamud said he hoped to have the new data base service operating by the end of the year.
Battle Brews Over Fees for SEC Reports

Group Believes Public Should Have Free Access

By Robert Thomason
Washington Post Staff Writer

N
eed electronic data about a publicly traded company? Traditionally, you've had to pay the information. Now, the government database that contains it is being tapped by on-line activists who, for the time being at least, are managing to distribute it for free.

It's all perfectly legal. With government and private funding, the Washington-based Internet Multicasting Service is buying raw financial data that companies file to the Securities and Exchange Commission's "Edgar" database, then putting it on the Internet computer network, where it can be retrieved by people worldwide.

The service has fueled a debate in many government agencies as cost-conscious Republicans come to power and electronic technology advances. Should agencies put on-line for free myriad information that they generate daily, or should they charge for it and save taxpayers some serious money?

"Pioneers of the information highway say the public deserves free, or at least very inexpensive, access to the SEC's huge Edgar system, which contains reports that publicly traded companies are required to file with the SEC."

But the SEC is resisting, choosing instead to sell the data to private companies that resell it and sell it on to the public. The SEC is required to fund itself, officials point out, and revenue from these sales is a big help.

The Internet Multicasting Service, a small nonprofit organization developing new uses of the Internet, began its SEC work in January 1993 with a grant from the National Science Foundation. That grant expires this year, so the service is lining up private help. So far, money or in-kind help has been pledged by Massachusetts Institute of Technology and New York University, Sun Microsystems Inc., MCI Communications Corp., RR Donnelley & Sons Co., and Time Inc.

Carl Malamud, president of Internet Multicasting, has called the collaboration the Information Highway Beautification Fund. Although his goal is the successful fund-raising effort, he still believes that the government should cover the cost of disseminating information that citizens have helped to generate by paying their taxes.

Internet Multicasting and its ally in this effort, Taxpayer Assets Project, are heartened by recent Republican moves to expand on-line access to government information. But they have not received commitments from GOP lawmakers about free public access to Edgar.

"The role of the SEC is to provide information to the market," Malamud said.

As the federal government expands its presence in electronic media, putting the data on the Internet for free would be an excellent way to pursue the SEC's goal, he said. He estimated that an Internet operation would require about $100,000 to start up and would need two employees to maintain.

But the SEC, which experienced years of delays and more than $20 million in cost overruns to launch the Edgar project in the first place, is loath to embark on new ventures when experienced and well-known companies pay well for the opportunity to do it.

David Copenhafer, who directs Edgar for the SEC, says the agency has neither the technical expertise nor the legal mandate to supply financial information in the sophisticated ways demanded by many computer users. The SEC collects raw data, but the data-service companies don't typically resell it that way; they "add value" by breaking it down into meaningful units.

"There will be hundreds of different products making use of the Edgar data to meet a tremendous spectrum of needs," Copenhafer said.

The database industry can meet these needs through their own technologies, he said, while the SEC cannot pretend to meet these needs with a single Internet database.

"The information already is becoming available through a variety of computer media." Pioneers of the information highway say.

Several companies are loading SEC data on CD-ROM discs. Moody's Investor Service Inc. is offering CD-ROMs of Edgar data at $995. Another company, Edgar Express Corp. of Florida, provides filings on a CD-ROM, but charges $4 for the password necessary to access the data in a single file in it.

In addition, the SEC is working with the Government Printing Office, which will produce a CD-ROM of Edgar data. And a local firm, DRT Design of Fairfax, is training companies to file to the new system.

Dayton, Ohio-based Lexis-Nexis is the prime SEC contractor that disseminates on-line data. The firm sells instantaneous feeds of SEC filings at a regulated annual prices of $183,000 or $138,000, depending on the speed of transmission used. It also provides, each day, computer tapes of the previous day's filings for $78,000 per year. The information is also available on Lexis-Nexis.

Sharon O'Donoghue, director of corporate legal markets for Lexis-Nexis, predicted that inexpensive Internet access to raw SEC data would not cut into Lexis-Nexis's business.

Most business customers do not want entire SEC documents, which are now available on the Internet and can be more than 100 pages long, O'Donoghue said.

Instead, she said, they want highly specific information about companies that would enhance their particular investment strategies. And some need it in a matter of seconds, or else the information is useless, she added.

In development since 1983, Edgar is scheduled to include filings from all of the 15,000 public companies this summer. The SEC operates the system in offices in Springfield, where it receives the data from the filing companies, compiles it and ships it electronically.

To access Internet Multicasting's Edgar data, Internet subscribers can use the gopher function to reach gopher.townhall.org, the ftp function to reach ftp.townhall.org, or the World Wide Web to reach www.townhall.org. Information can also be obtained through electronic mail by sending a message to mail@townhall.org, with the word help as the text of the message.
August 4, 1995

Dear Commissioner Lehman:

Since January, 1994, the non-profit Internet Multicasting Service has maintained an on-line database of U.S. Patent documents, as well as the full text of SEC Edgar documents, for the American public to access at no charge. This service will terminate on October 1, 1995.

Our on-line service allows keyword searches of the full text of U.S. Patents using the WAIS, World Wide Web, Gopher, and Electronic Mail services. Bulk transfers are available via the File Transfer Protocol. A digital signature is added to each document to allow users to verify the contents. Background material and links to additional resources are also maintained on the system.

Since January, 1994, we've sent out over 1.5 million Patent documents to the public. Over 2900 documents per day are now being distributed, and there is a strong demand by users for the full database (the more fully developed SEC system, for example, sends out over 17,000 documents per day). Our current database of patent documents is 180,000 files and 7.1 Gbytes. Given the current average rate of data, we estimate that a year of your data is approximately 4.5 Gbytes.

Based on the access figures and the size of your database, it would be a trivial task to put the entire Patent database available on the Internet. The cost of computer equipment, routers, and disk drives is well under $150,000. Our staff maintains the Patent dissemination service for well under 0.5 FTE of a systems programmer and 0.5 FTE of a developer, for a total labor cost of under $100,000. Internet access is under $25,000/year. If you depreciate the equipment over 3 years, the total cost per year of providing Internet access to your entire database is $175,000.

Our Patent and EDGAR service has been funded partially by a grant from the National Science Foundation. Additional support was provided by our corporate sponsors, by a coalition of university research centers, and with a great deal of "sweat equity" and our own money added by our dedicated staff committed to this important public service.

Recently, the Congress passed and the President signed the Paperwork Reduction Act, which becomes law October 1. This bill was a key provision of the Contract with the

THE INTERNET MULTICASTING SERVICE
America and has received the strong support of the Clinton administration. The law reads:

"(d) With respect to information dissemination, each agency shall--

"(I) ensure that the public has timely and equitable access to the agency's public information, including ensuring such access through--

"(A) encouraging a diversity of public and private sources for information based on government public information;

"(B) in cases in which the agency provides public information maintained in electronic format, providing timely and equitable access to the underlying data (in whole or in part); and

"(C) agency dissemination of public information in an efficient, effective, and economical manner;

In case there is any doubt, the committee reports go on to stress that information dissemination is an "integral part of the information life cycle," and emphasize that the purpose of the law is to "enunciate clearly the obligation of Federal agencies to ensure effective public access to government information."

Many of us have learned from the last two years of this demonstration project. It has been remarkable to see how large a portion of the American public has interest in these key documents. It has also been remarkable to see the advances in technology that allow dissemination to all members of the public a realistic and cost-effective goal. We hope you will embrace both the spirit and the letter of the laws recently passed by the Congress and signed by the President and that you will ensure that Patent data remains available on-line and does not get auctioned off to the highest bidder. We look forward to seeing your Internet service and will continue to be available to your staff as you make this important transition to the information age.

Sincerely,

Carl Malamud

cc: Vice President Gore
Hon. Ron Brown, Secretary of Commerce
Ms. Sally Katzen, Office of Management and Budget

Hon. Newt Gingrich, Speaker of the House
Hon. Robert Dole, Majority Leader of the Senate
Hon. Connie Mack, Republican Conference Secretary
Hon. Thomas Daschle, Minority Leader of the Senate
Hon. Richard Gephardt, Minority Leader of the House