

SiliconValley.com

INSIDE THE TECH ECONOMY

Mentor
Graphics

Advantage[®] **FPGA**

Our Advantage can be Your Advantage -
FPGA Advantage

[click here >>](#)

Search

Recent News

Archives

Web for

Welcome **Guest**
[Sign Up](#) | [Sign In](#) | [Member Benefits](#)

[Back to Home >](#)

Monday, Jun 20, 2005

News

- Mercury News
- Headlines

Columnists

Research

Venture Capital

Personal Technology

People

Valley News

Hot Topics

Careers

RSS Feeds

ONLINE EXTRAS

Events

Archives

Newsletters

Features

- [VC Survey](#)
- [Salary Survey](#)
- [SV 150](#)

XML

 email this

 print this

 reprint or license this

Posted on Wed, Aug. 11, 2004

Internet archivist has modest goal: Store everything

By Matt Marshall
Mercury News

Brewster Kahle, founder of San Francisco's Internet Archive, burns with a mission. He wants to ensure universal access to all human knowledge. And now he thinks that goal is within our grasp.

The emergence of cheap data storage technology has made what once seemed a pipe dream distinctly possible -- digitizing and storing the entire Web, the world's 100 million books, 2 or 3 million audio recordings and millions more software programs, TV shows and videos.

"Storing this is a no-brainer," Kahle said.

He's making what has been digitalized so far freely accessible at www.archive.org. And he's built an "Internet Bookmobile," a van that drives around the country downloading public-domain books from the archive via a satellite Net link. The van recently got kicked out of Walden Pond in Massachusetts for giving away Henry David Thoreau books and upsetting local book merchants. He has also taken the Bookmobile to places that really need it -- Uganda, Egypt, India -- printing out books for children at \$1 a piece.

And then there's the archive's newer offerings -- 15,000 music concerts and 300 feature films.

The archive recently hired an engineer to design an affordable Petabox, which stores a million gigabytes of data. The Petabox will be ready in the fall.

Brewster's goal: Store everything. "It is possible," he proclaimed last week at a conference at IBM's Almaden Research Center in San Jose. "It could be one of the greatest achievements of all time."

The Library of Congress houses about 28 million books, and he estimates he can scan and digitize each book for \$10 a piece. That would cost about \$280 million, or the equivalent of half the Library's annual budget.



FREE TECHNICAL EVENTS

TOPICS INCLUDE:
Simulation | Layout | Verification | Extraction

[View Online Today](#)

CLICK HERE

SV.COM SPONSORS

Visit [TerraServer.com](#), the leader in online imagery -- Satellite, Topo Maps, Aerial Photos and more.

Learn about the latest tools and trends in EDA. Download [industry-leading technical papers today!](#)

Want the latest verification solution? ModelSim Assertion-based Verification. [Read the Paper.](#)

FREE FPGA
Technical Papers.
[Download today!](#)

Free IC
Nanometer
Technical Papers
*~*Click Here*~*

Laser hair
removal -- [Laser
medicine
technology](#)

The Web is growing at about 20 terabytes of compressed data a month, which is manageable, Kahle said.

Sure, getting copyrighted material has its challenges, especially music and videos. But he's chipping away where he can.

Driving Kahle is the conviction that the world's information is a common good.

In that spirit, he also has asked Google to furnish him with a copy of its database, say with a six-month delay so Google's competitiveness doesn't suffer.

Google has yet to grant his request. But Kahle hopes the company will come around, especially in light of its claim that it wants to have a positive impact on the world. A Google spokeswoman declined to comment.

He learned a tough lesson when search engine Infoseek initially agreed to give him a copy of its database. When Infoseek went bankrupt, though, the lawyers didn't follow through. So Kahle is adamant that Google should act soon.

``Let's give it to others," he says.

Contact Matt Marshall at mmarshall@mercurynews.com or (415) 477-2518.



**REAL
Cities** Click here
to visit other
Real Cities sites

News | Columnists | Research | Venture Capital | Personal Technology | SV Life
About SiliconValley.com | About the Real Cities Network | Terms of Use & Privacy Statement | About Knight Ridder |
Copyright | RSS Feeds