



WEBFARMING

Newsletter - June, 1999

In This Issue

- The Web's Librarian
- Told You So!
- Google Gets Funding
- WF Seminars
- New WF Articles
- Tool of the Month: MagniFind
- XML in DB2
- Hyper-Searching the Web
- MiningCo is now About.com

R. Hackathorn
editor

[Comments?](#)

[Previous Issues](#)

[PDF Version](#)

[Home Page for WebFarming.com](#)

[The Web Farming Book](#)

[An Introduction to Web Farming](#)

The Web's Librarian

Brewster Kahle received some well-deserved recognition recently. His labor-of-love, **Alexa Internet**, was sold to **Amazon.com** for over \$250 million, which is quite a sum for a 50-person company with under a million in revenue and no profit. Alexa's core idea is to map the *foot-tracks* of people surfing the Web as clues for finding the better resources. The deal with Amazon includes continued support for the **Internet Archive**, a nonprofit that is preserving the Web for historical research.



[more ►](#)

Over ten years ago, Brewster invented the **WAIS** (Wide Area Information Server) system, pioneering Internet publishing.

WF Seminars

A three-day seminar on **Web Farming** is offered by **DCI** on the following dates:

- **September 15-17** in San Francisco
- **November 10-12** in Dallas

For full details, see the [online brochure](#). Seminars in Rome and Brussels are planned for October.

Tool of the Month

InXight is offering a free file viewer called **MagniFind**. It uses their Hyperbolic Tree visualization that allows quick navigation of complex hierarchical structures. It is a great way to try this visualization on your own local data. While visiting their site, also try the site map, which loads a Java applet using the Hyperbolic Tree.



Hyper-Searching the Web

Told You So!

Last week **Company Sleuth**, a service of **Infonautics**, created quite a stir over the acquisition of **SkyTel** by **MCI WorldCom**.

[more ►](#)

Google Gets Funding

Goodle, a new search service, received \$25M in venture funding from **Sequoia Capital** and **Kleiner**



Perkins Caufield & Byers for a 40% equity stake. CEO **Larry Page** (26 years old) and President **Sergey Brin** (25) started a year ago and have 23 employees. As reported in the **WSJ**, the site gets 800K visits per day, with a 50% monthly growth rate. Google emerges from a research project at **Stanford** and uses an algorithm similar to IBM's **Clever Project** (see below left).

New WF Articles

Two feature articles on Web Farming were published this month. **DM Review** contained the article "Farming Web Resources for the Data Warehouse" in their June issue.



DB2 Magazine printed the article "Web Farming" in their summer issue.



XML Support in DB2

As real progress in XML adoption, **IBM** is enhancing the text extender in the new **DB2 UDB** Version 6.1 to handle XML and other structured text. SQL searches can be limited to sections and even nested sections within an XML document. DB2 includes an XML parser so that XML documents can populate DB2 tables and DB2 tables can generate XML tagged documents. Entering beta, the new text extender will push XML support into the heart of the engine.

[more ►](#)

MiningCo is now About.com

In the latest Scientific American, an [article](#) explains the logic behind the **Clever Project**, work conducted at the IBM [Almaden Research Center](#). The logic is based on an iterative algorithm that classifies web pages as **hubs** (pages that link to many authorities) or **authorities** (pages that are linked to by many hubs). By detecting clusters of these pages for a topic, a **community** can be identified, as an aid in information searching. [more ▶](#)

In May, **MiningCo.com** changed their name to **About.com**. The site is noted for their expert guides on 650 topics, along with supporting discussion groups.

About.com



WEBFARMING

Newsletter - June, 1999

Told You So!

Company Sleuth, a free service of **Infonautics**, created quite a stir over the past week, amid headlines about a major corporate acquisition. Company Sleuth was reviewed in last February's issue of the WF Newsletter.

On May 25, Company Sleuth notified its clients who were sleuthing **MCI Worldcom** that the company had applied for a domain name of:

- SKYTELWORLD.COM

The newsgroups and message boards (such as, Yahoo! and Raging Bull) buzzed, causing the **SkyTel** stock to rise 7% on a volume over three times normal. As reported in the WSJ on the next day, MCI WorldCom executives responded by blaming 'an overzealous employee trying to protect the company from domain-name squatters.' They added that "the action is not an indication of any official company intention."

Guess what? Three days later, MCI Worldcom announced its intent to acquire SkyTel with a \$1.3 billion dollar stock deal.

It was a great publicity coup for Infonautics, achieving front page of the second section in the May 26 WSJ. Josh Kopelman, EVP at Infonautics, was quoted as saying that his service "scours the Web and finds clues." The WSJ continued with ". . . Internet addresses and other crumbs of data scattered around the Web are becoming important -- and often unintentional -- indicators of companies' business plans."

The story continues with the New York Times reporting on June 3 that some investors feel misled by MCI WorldCom statements with a possible liability for their losses.

The Infonautics people could not contain their excitement. In their Company Sleuth reports this week, they added at the top of their reports: **TOLD YOU SO!**

Of course, web-sleuthing works both ways. Last week, Infonautics secured the following domain names. Wonder what Infonautics is planning?

- PATENTSLEUTH.COM, REALESTATESLEUTH.COM, PRODUCTSLEUTH.COM, TRADEMARKSLEUTH.COM, Y2KSLEUTH.COM, TRAVEL-SLEUTH.COM, POLITICSSLEUTH.COM, RESEARCHSLEUTH.COM, NEWSLEUTH.COM, SUBMITJOBS.COM, RECRUITINGSLEUTH.COM, POLITICSLEUTH.COM, PERSONSLEUTH.COM, SHOPPINGSLAUGHT.COM

It will be increasingly difficult to avoid making tracks in webspace, as a company plans their new initiatives in their industry. This example is a solid endorsement of Web Farming as a critical component in future business intelligence systems.

- Richard Hackathorn