

The Big Picture: In Focus

The following is excerpted from the March 2004 issue of Release 1.0.

Technorati: We see the conversation

BY CHRISTINA KOUKKOS

“I founded Technorati because I wanted to know who was talking about me,” says CEO and founder Dave Sifry, who was previously CTO and founder of Linuxcare and CTO and founder of WiFi technology provider Sputnik. “The infrastructure of the Web is changing. An increasing number of people are using new tools to create content and publish it to the Web - weblogging software, camera phones - causing an explosion of user-generated content. The business opportunity is to aggregate this and provide back a mirror of what’s going on in real-time - around individuals, or companies, or competitors. . . whatever someone might want to explore or track.”

In essence, Technorati takes Google’s fascination with links into the new world of the blogosphere and opens up a lot of specific information beyond just “search relevance.” It allows users to see which bloggers are linking to a particular Web page – its “cosmos” – as well as to search by keyword or URL, with results ranked either by time (the most recently published page that mentions the subject or links to the URL) or “blog authority” (the most linked-to blogs or websites that mention the subject or link to the URL).

“Try to find out anything on Google or Yahoo! that happened in the last two weeks and that’s not breaking news,” says Sifry. “You’re lucky if you pick up a PR release. But we’re built into the real-time blogging infrastructure.” Technorati has published an open API and persuaded the major blogging tool vendors to use it; blogs published with those tools automatically ping Technorati each time they are updated. Technorati gets about 150,000 pings per day. This API is something any search engine could adopt, but Technorati has built relationships with publishers over time and now covers many more blogs than its competitors – about 2 million, with 100 million active links. For companies such as Google and Yahoo!, he contends, “adding such a capability would be an unnatural act; they are set up for a different model.”

{ continued on page 2 }



The conversation starts here.

Release 1.0 is published 11 times per year. An annual subscription is \$795 (\$850 outside the US, Canada and Mexico) and a single issue is \$80. No material in this publication may be reproduced, displayed, modified, copied, published or sent to another party, in whole or in part, without the prior written consent of CNET Networks, Inc. in each instance. Copyright(c) 2004, CNET Networks, Inc.

Release 1.0® (ISSN 1047-935X) is published monthly except for a combined July/August issue by EDventure Holdings Inc., 104 Fifth Avenue, New York, NY 10011-6987; 1 (212) 924-8800; fax, 1 (212) 924-0240; www.edventure.com. It covers the worlds of information technology and the Internet, including wireless communications, security, business models, online services, tracking systems, identity management and other unpredictable topics. . . and the policy issues they raise.

EDITOR: Esther Dyson
(edyson@edventure.com)

PUBLISHER: Daphne Kis
(daphne@edventure.com)

MANAGING EDITOR: Christina Koukkos
(christina@edventure.com)

CONTRIBUTING WRITERS: Dan Gillmor
(dan@gillmor.com),
John Hagel (john@johnhagel.com),
Steven Johnson (stevenberlin-johnson@earthlink.net),
Clay Shirky (clay@shirky.com),
Jeff Ubois (jeff@ubois.com),
Dave Weinberger
(self@evident.com)

CIRCULATION MANAGER: Brodie Crawford
(brodie@edventure.com)

SYSTEMS MANAGER: Geoff Clarke
(geoff@edventure.com)

CONSULTING EDITOR: Bill Kutik
(bill@kutik.com)

Copyright © 2004, EDventure Holdings Inc. All rights reserved. No material in this publication may be reproduced without written permission; however, we gladly arrange for reprints, bulk orders or site licenses. Subscriptions cost \$795 per year in the US, Canada and Mexico; \$850 overseas.

Sifry continues, “The Googles and Inktomis do documents. We recognize that documents come from people. So we look at the social aspects: blogrolls, who quotes whom, who has influence. Because we understand time, we can watch how influence spreads and how opinion focuses, changes and moves on. We see the blogosphere as a giant threaded conversation, and we are discovering and tracking the thought leaders and early adopters.”

Once you have such basic data, you can find out lots of interesting things about the “shape” of the overall conversation or of any particular cluster: The service publishes lists such as the top 100 websites linked to by blogs, interesting newcomers, the products (mostly books) and news stories that are generating the most buzz among bloggers, and so on. Users can also set up watchlists to track commentary by keywords or URLs – a buzzmeter, basically.

One interesting fact about the blogosphere that Technorati visualizes is its long tail. In most endeavors, the top end of the market – say, blogs with 1000 inbound links – would take the majority of the market share. But in the case of blogs, because they are so cheap and easy to publish, the tail end of the market accounts for a huge amount of fragmented activity. In fact, notes Sifry, citing figures from the site: “There are only 6,905 weblogs with inbound links from 50 or more unique blogs. There are 15,872 with 25 or more and 97,854 weblogs with 5 or more. There are 1.4 million blogs with no inbound links at all.” Technorati measures the stats this way to avoid overcounting blog-link spam.

Technorati is still relatively small, operating on 24 Linux servers, but it is optimized to manage a seven-minute latency. It adds 12,000 new blogs a day (and sees about a third of the new ones wither away after three months).

Technorati’s business model is to sell such watchlist services, with various enhancements, to marketers and PR firms, Wall Street and other market analysts and others for whom real-time market information is money.