



Google's Dataspaces Technology Dings PageRank

October 31, 2007

By Clint Boulton

The drops in Google PageRank ratings that plagued the blogosphere last week are a result of a data management abstraction layer Google has implemented to improve its search algorithm, according to an IT expert and author.

ADVERTISEMENT

Safecount.net

TAKE A SURVEY

Research purposes only.

Sites such as WashingtonPost.com, Forbes.com and EnGadget.com, which have seen their PageRank scores fall, are seeing the first fruits of dataspaces.

- RELATED LINKS**
- [Google Mini Goes for the Max](#)
 - [Google Product Chief: It's Okay to Fail Wisely](#)
 - [What's Your Google Rep?](#)
 - [EMC Offers Google Desktop Search to Enterprises](#)
 - [Google Gets Enterprising](#)

Dataspaces is a technology incorporated in Google's search algorithm, which includes filters that sift out attempts by site owners to boost their PageRank rating against Google's Webmaster guidelines, said Stephen Arnold, who studies Google's patents and technical documents.

The technology, which is being implemented by Alon Halevy, a Google engineer and former University of Washington computer science professor with an extensive background in database mining and business intelligence, builds on Google's programmable search engine methodology.

"What people are reacting to is the first of a series of waves of Google technology designed to eliminate abuse, increase the type of ads that are available and, of course, Google's monetization," Arnold told eWEEK Oct. 26.

[Click here to read more about sites railing about PageRank decreases.](#)

A Google spokesperson denied the connection, noting that dataspaces had nothing to do with search or PageRank. Google, secretive about works in progress, also declined to provide more information on how it is using dataspaces.

Arnold said that Google is using dataspaces to put different content in a single, relevance-ranked results list, and said as much in a paper, which he delivered at the Infonortics ICIC 2007 show in Spain Oct. 22.

"The idea is that instead of a single list of results from Web content or Google Books, the results list would present many types of content, each relevant to the query, and instantly accessible from the Google interface," Arnold [wrote in his paper](#), which concluded that Google has a solid foundation for data and text mining.

Google engineer Alon Halevy is the driving force behind dataspaces, and has written or co-authored several papers on the subject, including "[Indexing with Dataspaces](#)," and "[Dataspaces: The Next Frontier for Data Integration](#)," which he delivered as part of a lecture series at San Jose State University this year.

While dataspaces seem like a technology approach suited for enterprises, Halevy believes dataspaces should be applied to consumer-oriented applications, an approach that plays to Google's specialty.

[Click here to read more about Google boosting its analytics tools.](#)

"I believe that the data management community should shift its focus away from enterprise computing and consider consumer-facing applications," Halevy [wrote on his Web page](#) for the University of Washington.

Halevy created software maker Nimble Technology, which he sold to Actuate in 2003. He then founded Transformic, which developed an engine for searching databases that reside behind Web sites, and sold it to Google in 2006.

As much as dataspaces helps people find content, Arnold said it will help weed out sites where people are trying to "game" Google's ranking system through link farms and paid listings. Arnold believes his research highlights that Google is serious about providing search results uncompromised by scheming SEO (search engine optimization) efforts.

"People that squeal are behind the curve because their shortcuts are being caught out," Arnold said, scoffing at the suggestion that Google is deliberately tweaking its PageRank system to mess with sites.

Arnold said the search algorithm is all math all of the time for Google, noting that if site operators followed the Webmaster guidelines set forth by Google, they wouldn't get hit with a lousy PageRank.

Arnold also predicted further tightening of the search algorithm via dataspaces in the coming weeks.

Check out eWEEK.com's Search Center for the latest news, views and analysis on enterprise search technology.

[Copyright \(c\) 2007Ziff Davis Enterprise Inc. All Rights Reserved.](#)