



Google Favored by Web Admins

Web-site policy makers are playing favorites, and Google is the beneficiary, researchers say.

Network World Staff

Thursday, November 15, 2007 05:00 PM PST

Web-site policy makers are playing favorites, and Google is the big beneficiary, say Penn State researchers.

The research team created a search engine called [BotSeer](#) and used it to examine more than 7,500 Web sites. The team found a pro-[Google](#) bias in terms of which search-engine Web crawlers were or were not allowed access.

Web site administrators use robots.txt files to regulate Web crawlers -- also known as spiders and bots -- and to help prevent servers from getting overloaded. The researchers found that about four in 10 sites used robots.txt files, up from one in 10 in 1996.

"We expected that robots.txt files would treat all search engines equally, or maybe disfavor certain obnoxious bots, so we were surprised to discover a strong correlation between the robots favored and the search engines' market share," said C. Lee Giles, the David Reese Professor of Information Sciences and Technology at Penn State who led the research team that developed BotSeer, in a statement.

"Robots.txt files are written by Web policy-makers and administrators who have to intentionally specify Google as the favored search engine," Giles said.

[Yahoo](#)

and MSN also were given greater-than-average access to Web sites, but Google showed up in robots.txt files more than twice as often as either of these competing search sites.

The findings are [described in greater detail](#) in the paper "Determining Bias to Search Engines from Robots.txt."

Other recent network-oriented Penn State research includes efforts to advance [photo searching](#), push [copper network links](#) to support higher data rates, and [better secure databases](#).

For the latest on network-oriented research at university and other labs, go to Network World's [Alpha Doggs](#) blog.

A promotional graphic for the "FREE Rice" campaign. It features a green and white color scheme. At the top left, the word "FREE" is in small letters above "Rice" in large letters. To the right is a photo of a young child eating rice. Further right is the WFP (United Nations World Food Programme) logo. Below the photo is a green box with a white arrow pointing right and the text "play the internet craze and feed the world's hungry at the same time!". At the bottom is another green box with white text: "For each word you get right, 20 grains of rice is donated to the World Food Programme to help end hunger - more than 15 billion grains to date!".