

Web Watcher

BY GLENN FLEISHMAN



References Available Upon Request

How VISITORS GET TO YOUR SITE may seem less important than *how many* folks get there, but understanding the where-from can help focus your what-to-do.

Unless you're a heavy traditional media advertiser, a newspaper, or a magazine, most people reaching your site probably aren't getting there by typing in a Web address. Furthermore, unless you're buying Web banner ads out your ears, you're probably not getting many visitors from online ads, either. Visitors are finding you by clicking on links at other sites and search engines. So your job is to find out what they're clicking—what links from what URLs—and use that information to improve traffic to your site.

How do you measure where your visitors come from? Every major Web-server software package—like Apache or Microsoft's Internet Information Server (IIS)—allows logging of what are called "referring URLs," the URLs from which visitors click to get to your site. Most browsers automatically send their referring URL when they request a page

from a server. Originally, this helped promote better linkage of disparate information. Now it's a marketing treasure trove.

About 40 to 60 percent of all visitors to a Web site reveal their referring URLs. Whether or not visitors reveal their referring URLs depends on what browser they use and whether they're surfing from behind corporate firewalls—security systems that function as watchful intermediaries between companies' intranets and the Internet—which often block referring URLs. The browsers from Netscape and Microsoft pass referring URLs without comment, even though both offer significant control over cookies—the bits of data that Web sites can store and read on users' computers to track surfing behavior.

Since you'll be able to gather only partial data on referring URLs, it's important to use that data as guidance, not gospel. For example, because AOL—remarkably—doesn't send referring URLs, you'll be missing information on an entire consumer segment.

Accessing this information for analysis should be straightforward. First, your Web server needs to be configured to keep a record of referring URLs in a Web log—a text file to which information is appended a line at a time, one line for each page or graphic requested.

Analyzing referring URLs requires either a programmer or log-analysis software. A programmer

can write some simple code that walks through the Web log and summarizes its information. And there are many software packages, ranging in price from a few hundred to tens of thousands of dollars, that can read Web logs and perform an astounding variety of analyses, including reports on referring URLs.

The analysis should show you trends. Many sites get significant traffic from search engines, like Excite, and from directory services, like Yahoo. If you feel you're not getting enough traffic from these sites, refer to my column "Getting to the Top" (Autumn 1998 issue), which explains how you can tweak your pages so that search engines are more likely to list you higher in search results lists. Later, you can analyze referring URLs to find out whether these changes actually affected traffic.

Another strategy for increasing site traffic is through affiliate programs or similar "bounty" relationships, in which companies pay an associated site for incoming traffic that results in page views, sales, or hot leads. You can examine logs to figure out what sites are sending clients your way so that you can investigate setting up more formal affiliate relationships with them.

Then you can use referring URL data to get some sense of how successful these programs are (just be careful not to take this data at face value—it's too unreliable to track the effectiveness of these relationships very accurately). But the more you know about how people reach your site, the better you can plot how to extend your own reach. ♦

Glenn Fleishman is a contributing editor for Adobe Magazine and has written articles for The New York Times "Circuits" section. His latest project in information technology is the design of <http://isbn.nu>.