

# Web Watcher



By Glenn Fleishman

## Visits, visitors, and hits

**A particularly confusing** part of reading about the Web is trying to determine how many people actually come to a Web site. Commonly, the term “hits” is used, but that’s a wildly inaccurate term to present in raw form. What companies running Web sites, potential advertisers, and—increasingly—knowledgeable readers want to know is: How many visits have there been? How many unique people are there? What are their demographics? And what do they do when they visit?

Let’s clarify the terminology. Though you’ll see variations on the scheme that follows, it is gradually being accepted by the computer industry as standard usage.

**Hit:** A request by a browser for a file, or, technically, a user’s click on a link and the resultant file transfers initiated by the browser. An HTML page usually has several images on it; when a user enters a URL or clicks on a link, the browser first retrieves the HTML page itself (the formatted text) and then scans it for files it has to retrieve to display on the page. Separate requests for those images are initiated simultaneous to the display of the text portion of the page. A typical page, like Adobe Systems’s home page, might result in eight or nine separate requests—one for the HTML page and the rest for images. Each of these requests is a “hit.”

**Visit:** A unique browser from a unique location generating requests on a site (browsing or surfing) over a discrete period of time. Typically, Web producers bound visits on either end, so that if a specific location and browser doesn’t show up in their log of requests for, say, 30 minutes, the next appearance is a unique visit.

**Visitor:** A unique user. Identifying users is done in several ways that vary from simply using a cookie (a tag that can be assigned to the user; for more, see the next item) to requiring user registration. If you know the visitor, whether by name or by cookie, you can do further analysis to

know how many times that visitor comes to the site, where he or she goes, and so forth. This makes some people’s skin crawl; these users would prefer to be anonymous. But some Web sites use this kind of information—with the user’s permission—to create profiles and custom-tailor information and services to them.

**Cookie:** A persistent bit of information, stored on the user’s local hard drive, that is keyed to a specific server (and even a file pathway or directory location at the server) and passed back to the server as part of the negotiation that takes place when the user’s browser again crosses that specific server/path combination. Cookies have been portrayed as invasions of privacy, but the truth is that you can’t use a cookie to retrieve information *about* a user. You can use it only to record whether someone previously visited a site, and what that particular user’s viewing behavior is. This is done by storing small codes or tokens that can be arbitrary (a sequential or random number that’s unique) or mapped to information provided by the user. The Navigator browser, according to Netscape’s specifications for it, will store only 4,096 bytes in a cookie, and only 300 cookies overall (a maximum of only about 20 per site), before dumping the information in first-in, first-out fashion.

Many reporters get all this wrong, which is why you read statements like “More than a million people visited *umathurman.com* on its first day.” In reality, the number of visitors is often a factor of 20 to 40 smaller—the reporter has undoubtedly confused “people” or visitors with hits, and probably confused visits and visitors as well. A million hits could translate to 20,000 visits (and you won’t know how many of those visits are by people returning once or more in the same day), based on an average of 50 hits per visit.

You might also see the terms “impression” and “cost per thousand” thrown

about. The number of impressions is the number of unique times an advertisement is displayed to a user. In the print-magazine world, this often includes the pass-along rate—not just the number of subscribers, but an audited number showing how many people, on average, pick up and read an issue and thus (presumably) see a given ad. In the Web world, the number of impressions can be measured exactly, but the number of unique individuals cannot, although cookie-tracking can help.

Cost per thousand—abbreviated CPM, with the M being the Roman numeral for thousand—represents the unit in which ad dollars are allotted. A typical CPM for Web advertising (and for print magazine advertising as well) is \$30. That is, for every 1,000 impressions, a site would get \$30 (less commissions and other details) from the advertiser.

**A final note:** Many sites are prone to exaggeration, and that’s partly because there’s no Nielsen or Arbitron of the Net—yet! Several companies are hoping to position themselves as independent auditors to confirm for advertisers what the actual traffic is. One of them, Internet Profiles (I/PRO), is now part of Nielsen, while the Audit Bureau of Circulations, which conducts circulation audits for print magazines, has launched the Audit Bureau of Verification Services, Inc., which it hopes will do the same for Web sites.

While institutions develop to legitimize visitation claims, view the numbers that you see with skepticism. It’s likely that, after reading this short column, you know more about how the Web is spun than many of the folks writing about it. ▀

*Contributing editor Glenn Fleishman, a principal at Point of Presence Company, can be reached at [glenn@popco.com](mailto:glenn@popco.com). His column “Web Watcher” appears in each issue of Adobe Magazine, and he also contributes to [adobe.mag](http://www.adobemag.com) (<http://www.adobemag.com>).*



Contents of this file and the entire contents of this CD © 1996 Adobe Systems Incorporated, all rights reserved.

Adobe, the Adobe logo, Acrobat, Acrobat Capture, Acrobat Catalog, Acrobat Exchange, the Acrobat logo, Acrobat Reader, Acrobat Search, Adobe Caslon, Adobe Dimensions, Adobe Font Folio, Adobe Garamond, Adobe Illustrator, Adobe Jenson, Adobe Magazine, Adobe Originals, Adobe Premiere, Adobe Table, Adobe Type Manager, Adobe Wood Type, After Effects, Aldus, Aldus Magazine, Art Explorer, ATM, Birch, Blackoak, Caflich Script, Charlemagne, Checklist, Classroom in a Book, Classroom in a Box, CoSA, Cottonwood, CustomerFirst, Distiller, Ex Ponto, FaxYI, Frame, FrameMaker, FrameViewer, Galahad, Gallery Effects, Giddyap, Giddyap Thangs, HomePublisher, Ironwood, Jimbo, Lithos, Madrone, Mesquite, Mezz, Minion, Myriad, Nueva, PageMaker, PageMill, Penumbra, Persuasion, PhotoDeluxe, Photoshop, Poetica, Ponderosa, Poplar, PostScript, PSpriinter, Sanvito, ScreenReady, SiteMill, Streamline, SuperATM, SuperPaint, Tekton, TextureMaker, TIFF, Trajan, TypeAlign, Type On Call, Type Reunion, Type Twister, Viva, and WildType are trademarks of Adobe Systems Incorporated or its subsidiaries and may be registered in certain jurisdictions. InfoPublisher is a registered trademark of Page Ahead Software, licensed to Adobe. Macintosh and Power Macintosh are registered trademarks of Apple Computer Inc. Windows is a registered trademark of Microsoft Corporation. Other products and corporate names may be trademarks of other companies, and are used only for explanation and to the owners' benefit, without intent to infringe.

All product pricing and offers represented in *Adobe Magazine* are applicable in the U.S. and Canada only. Prices and availability outside this area may vary; residents of other countries should check with their distributors regarding offers available to them.

**browse**