To determine the prevalence of Internet tracking technologies, The Wall Street Journal analyzed the 50 most-visited U.S. websites, as ranked by the comScore Media Metrix report from October 2009.

The Journal hired a technology consultant, Ashkan Soltani, to analyze the 50 sites for three types of tracking methods commonly used online: "HTML cookies," "Flash cookies" and "beacons."

HTML cookies are small text files, installed on a user's computer by a website, that assign the user's computer a unique identity and can track the user's movements on a site. Flash cookies are used in conjunction with Adobe Systems' Flash software, which is widely used to display graphics and video on websites. Beacons are bits of software code on a site that can transmit data about a user's browsing behavior.

Mr. Soltani visited the 50 sites between Dec. 10 and Jan. 14. Before each session, he cleared his computer of all browser data, including HTML cookies, Flash cookies and beacons. Each session consisted of visiting 20 pages per site. In one case, involving PayPal, he visited only six pages because viewing more would have required logging in to the PayPal service.

Mr. Soltani used Mozilla Firefox 3.5 and Adobe Flash Player 10.0. Following each session, he examined the tracking files that had been placed on the computer.

Beacons typically don't place a file on a computer. To trace them, Mr. Soltani used Ghostery, a small piece of software that can tell if a beacon is sending information from the website being examined.
Mr. Soltani also used a network-analyzer program to record all communication during a session, and to identify when his computer connected to other sites, to download an ad, for example.

At the time of his hiring by the Journal, Mr. Soltani was an independent consultant. For his master's thesis at the University of California, Berkeley, he and co-authors analyzed the use of beacons at the top 100 U.S. websites. He is now a contract technologist at the Federal Trade Commission. The FTC had no role in this study.

The Journal database also contains information collected by PrivacyChoice LLC about the privacy policies of companies that place these tracking files on websites. PrivacyChoice, founded by tech entrepreneur Jim Brock, provides privacy-consulting services to websites and doesn't accept money from ad companies that it surveys.

PrivacyChoice also provided the technology for the TrackerScan software that The Wall Street Journal is offering to readers to determine what cookies and other tracking tools are present on their own computers. You can access the software at WSJ.com/wtk.

The Journal compiled an "exposure index" for the 50 sites it examined, combining Mr. Soltani’s findings with PrivacyChoice's analysis of cookie-placers, to determine how much each site exposes visitors to intrusive monitoring.

The exposure index gives each site a score based on eight criteria in PrivacyChoice's analysis: whether the site belongs to an industry self-regulating group; whether it lets users opt out of receiving cookies; whether it is part of an advertising or tracking network; whether it shares data it collects with others; whether it promises to keep user data anonymous; how long it retains user data; and how it handles sensitive data such as financial or health information.

A site’s exposure index is the sum of the scores for each cookie, beacon and Flash cookie found on that site. The Journal used statistical analysis to group the 50 sites into four clusters of sites with generally similar characteristics.