

Web Watcher

BY GLENN FLEISHMAN



Nowhere to Hide

I'M CURRENTLY BEING PUBLISHED IN Italian without permission. How did I find out? Through "ego surfing"—I searched on my name as a phrase at some popular search engines and found an unauthorized reprint.

The work being plagiarized is an *Adobe Magazine* column entitled "Hits, Visits, and Visitors" (it was originally published in the September/October 1996 North American edition and as "Visite, Visitatori e Accessi" in the 2/1997 Italian edition). Some enterprising soul with no knowledge of copyright law—or with a pretension of ignorance—typed it in, scanned my photo even, and put it up. (Despite a few attempts, we still haven't gotten it removed or paid for, although they did take down the photo.)

Before the Internet, it was easy for thieves to lift the work of a writer, photographer, or illustrator and be pretty sure they'd never be found out, especially if the source was regional and the thief was reproducing it somewhere else. But the Internet is changing that. While the Web may make it easier to plagiarize others' work, it also provides tools that make it easier to track down cases of plagiarism.

The Net makes just about everything easy to search for and find. Although a very few large media sites—*The New York Times* and *The Wall Street Journal*, for instance—require registration and even payment to search their databases, most Internet content is indexed via

search engines and easily reachable through a few keystrokes.

One less-known search technique is to use diction, or word choice. First, think of a few uncommon words used in whatever written material you want to search for (to find this column, I might choose diction, visitatori, and ego). Then enter these words in any search engine, either by using a plus sign to force a match on all of the words or by selecting Match All Words from a pop-up menu, depending on the site. If the material you're searching for is on an indexed Web page, you'll find it—and any other page that reproduces your work or uses the same slightly obscure words.

Because diction is so idiosyncratic, it can be used strategically to protect written works. For instance, some corporations and government agencies create different versions of sensitive documents, and in each version they'll replace fairly unimportant words with synonyms. By tracking which synonyms were distributed to which groups of people, leaks to the media or other companies can be tracked in direct quotes or even paraphrases.

But what if you're looking to protect your images or illustrations? *The New Yorker* magazine faced this problem while preparing to launch the new Web site for its subsidiary The Cartoon Bank, which handles the searches and licensing for reprint rights to cartoons that appear in *The New Yorker*.

Bob Mankoff, the cartoons editor at *The New Yorker*, decided to use Digimarc's watermarking system to help track unauthorized use of cartoons. Digimarc's system (and similar ones from companies such as Signum Technologies) add to images carefully calculated noise that's virtually invisible to the human eye, but that can be detected with the watermarking software. The Cartoon Bank will also use a Digimarc service that trolls the Net, retrieves images, compares any watermarks it finds against its database, and then generates automated reports for the copyright holders who subscribe, letting them know where the work has ended up.

There's no absolute way to protect your work from plagiarism, especially if you post in HTML. (If you post in Adobe PDF—portable document format—you can set options that make it very hard for others to copy, modify, or even print documents without consent.) You need to have faith that most people will respect your property, and you can use search techniques and watermarking to help catch unauthorized use. As Ronald Reagan said about negotiating treaties with the then-Soviet Union, "Trust but verify." ♦

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