

## USING ELECTRONIC RESEARCH TO DETECT SOURCES OF PLAGIARIZED MATERIALS

BY ANNA M. CHERRY

*Anna M. Cherry is a Reference Librarian at William Mitchell College of Law in St. Paul, Minn.*

Plagiarism is a difficult issue for law schools, but not one that can be ignored. The problems of inconsistent definitions, policies, and enforcement among law schools have been recently discussed in an article aimed at students.<sup>1</sup> The subject also has been covered in detail in the academic literature.<sup>2</sup> But for the professor faced with a questionable paper, the first step is investigation.

The suspicion of plagiarism usually arises when a paper seems inconsistent with a student's other work through vocabulary, sentence structure, or style. Perhaps the professor finds the language in the paper reminiscent of something she read recently. A check of the footnoted materials may reveal plagiarism through carelessness or laziness. But what if it's not that easy?

Electronic research services and the Internet can be used to help identify copied sources through full-text Boolean searching. This article suggests ways to use LEXIS®, Westlaw®, and Internet search engines to identify language in published sources that matches language in a submitted writing. Also included are two Internet services that, for a fee, will check papers against texts received from term paper mills and documents on the Internet.

### LEXIS and Westlaw Searching

Westlaw and LEXIS have large, full-text databases that are both easy to copy from and easy to search for copied material. Choose Terms and Connectors searching as the search option.

Natural Language searching is not helpful in these circumstances because it retrieves results based on the relative weighting of the search words rather than using an exact word match.

For search terms, choose words and phrases from the suspicious paper that sound likely to have been copied. Choose longer and unique phrases, but don't choose something too long. The longer the phrase the more likely it is that some of the words will have been changed and no exact matches will be returned.

There are two strategies for making a search out of the phrase. The simplest is to put the entire phrase in quotation marks. This works better for shorter phrases. The other is to use a within-same-sentence connector—the /s—to connect the words within the phrase that are most likely to have been copied, eliminating words that may have been changed.

Here is an example paragraph copied from a recent law review article:

The procedural vacuum left by the statute and two decades of creative interpretation of it by various Minnesota appellate panels has created a perplexing situation for lawyers and judges faced with assessing whether a modification request should or will proceed. The uncertainty makes it difficult at this critical juncture of a modification dispute to predict with any degree of accuracy the weight a court will give to competing information before it. Furthermore, because trial judges have not received clear procedural guidance from the appellate courts, they may render inconsistent rulings.<sup>3</sup>

The following are examples of searches that will retrieve this article from Westlaw or LEXIS:

- “procedural vacuum” /s “creative interpretation” /s “perplexing situation”
- “uncertainty makes it difficult at this critical juncture”
- “trial judges” /s “clear procedural guidance” /s “inconsistent rulings”

The third search takes longer for the system to

<sup>1</sup> Lynn Weisberg, *Giving Credit Where It's Due*, Student Lawyer, Nov. 2000, at 19.

<sup>2</sup> Terri LeClercq, *Failure to Teach: Due Process and Law School Plagiarism*, 49 J. Legal Educ. 236 (1999).

<sup>3</sup> Robert E. Oliphant, *Redefining a Statute Out of Existence: Minnesota's View of When a Custody Modification Hearing Can Be Held*, 26 Wm. Mitchell L. R. 711, 712 (2000).

process because of the common words it uses, but all three retrieve only the one article.

For papers that have been heavily paraphrased, another suggestion is to search the databases using the citations from the questioned paper as search terms.<sup>4</sup> If an article contains the same citations as a student's paper, it may have formed the sum total of the student's research.

The principle of searching for citations in a text is the same as searching for a phrase. For example, the search "**21 wm. mitchell l. rev. 989**", including the quotation marks, will retrieve documents that have cited another article by Prof. Oliphant, "*Is Sweeping Change Possible? Minnesota Adopts the Uniform Interstate Family Support Act*," 21 Wm. Mitchell L. Rev. 989 (1996). Words from the citation can also be strung together with connectors to shorten typing or avoid missed retrievals due to typographical errors in the database. For example, **oliphant & uniform & 21 +s 989** will also retrieve citations to Professor Oliphant's article on Westlaw. On LEXIS, the equivalent search would be **oliphant pre/20 21 pre/5 989**. Both the +s and pre/n connectors require words to be in proximity and in the same order as entered. The difference is that Westlaw allows the option of having the words be in the same sentence whereas LEXIS requires the searcher to count the number of words between.

### LEXIS and Westlaw Databases

LEXIS and Westlaw have individual databases for individual titles and also combined databases lumping journals and other secondary sources together. Some of these combined databases are listed below. Very recent or unpublished cases should also not be overlooked as possible sources of copied materials. The online and printed database directories provide additional options. Calling customer service is another way to find out if a particular publication is online. The West Group Reference Attorneys can be reached at 1-800-REF-ATTY (1-800-733-2889). LEXIS customer service can be reached at 1-800-543-6862.

To find which individual titles are included in

<sup>4</sup> David J. Shakow, *Computers and Plagiarism*, 42 J. of Legal Educ. 458, 459 (1992).

these databases and the range of dates covered, check the scope or guide pages. When using the Web interface, information about a database can be found by clicking a button with the international information symbol (a small *i* in a circle) next to the database name in the directory. It is important to remember that these services were invented in the late 1970s. Texts and articles published before the early 1980s are probably not included.

### Some useful databases on Westlaw

**JLR**—The full text of hundreds of law reviews, CLE courses, and bar journals. Full coverage of all articles in the journals starts with 1994. Prior to that date, only selected coverage is available.

**TEXTS**—Texts, treatises, and legal encyclopedias in full text.

**ALR**—*American Law Reports* series, except ALR® First.

**TP-ALL**—The catch-all database for texts and law reviews.

**ALLNEWS**—Newspapers and other news sources. Note that an automatic date restriction limits searching to the most recent articles, unless the searcher specifies otherwise.

**MAGSPLUS**—Magazines, journals, and newsletters. Note that an automatic date restriction limits searching to the most recent articles, unless the searcher specifies otherwise.

### Some useful databases on LEXIS

LEXIS also includes hundreds of full-text law review articles; selected treatises; the *American Law Reports* series, except ALR First; and probably the world's largest database of magazines, newspapers, and other news sources. In the Web version of LEXIS, these files can be browsed by category. Look for law reviews and texts under *Secondary Legal*; look for news sources under *News*.

### Internet Search Engines

There is no single search engine for the World Wide Web. Different search engines work differently and have different strengths and weaknesses. Which ones have the most coverage also changes over time. Meta tag search engines that send the same query to multiple search engines are probably not a good choice because

they usually return only the top hits from each engine, often the same hits.

Unique words or short phrases can be searched anywhere, but the FindSame search engine—[www.findsame.com](http://www.findsame.com)—is the only one designed to match text by phrases, sentences, and paragraphs. A large chunk of text or an entire document is entered and the search engine looks for Web pages that contain significant fragments of that text.

Google—[www.google.com](http://www.google.com)—is a search engine that ranks its results not by how many keywords from your search match a page, but by how many other sites link to it. This makes Google especially useful for finding official Web sites for organizations—for example, a government agency or association that may have authored a report.

Sites such as Search Engine Showdown—[www.notess.com/search/](http://www.notess.com/search/)— and Search Engine Watch—[searchenginewatch.com/](http://searchenginewatch.com/)—evaluate search engines. Visit these sites to select a search engine and to compare features. Each search engine has its own quirks of operation. Most will accept quotation marks to denote a phrase. Most use a plus sign (+) to mark a word as required, e.g., +**hearing**. With any search engine, it is wise to read the search tips section before attempting the kind of complicated searching described in this article.

### Other Internet Resources

The Center for Excellence in Teaching and Learning of the University at Albany has created a useful Web page for teachers at [www.albany.edu/cetl/resources/plagiarism.html](http://www.albany.edu/cetl/resources/plagiarism.html).

New fee-based services aimed primarily at the college market offer testing of papers against databases of previously submitted papers, many from Internet term-paper mills, and searches of the Web. It should not be assumed that term-paper mills do not include law and law-related topics. Two of these services are [www.plagiarism.org](http://www.plagiarism.org) and [www.integriguard.com](http://www.integriguard.com). The latter site includes a free service for checking originality of papers at [www.howoriginal.com](http://www.howoriginal.com).

### Limitations of Online Research

All of the text-matching strategies discussed so far depend on the full text of a document being available online, but even in this electronic age, many are not. LegalTrac, for example, indexes hundreds of law review articles from 1980 to the present that are not available in full text on LEXIS or Westlaw. LegalTrac and Index to Legal Periodicals (from 1981) are available for searching through a Web subscription or through their LEXIS and Westlaw files.

Another possible resource not to overlook are continuing legal education materials. Many American Law Institute–American Bar Association (ALI-ABA) and Practising Law Institute (PLI) course materials are online on LEXIS or Westlaw. Local CLE materials should be searchable through a local law library's online catalog.

### Conclusion

In this age of online research, documents available from LEXIS, Westlaw, and the Internet make intentional and unintentional plagiarism easier. With a little creativity and persistence, the professor can use these tools to make detection of plagiarism easier too.

© 2001 Anna M. Cherry

“All of the text-matching strategies discussed so far depend on the full text of a document being available online, but even in this electronic age, many are not.”