THE MEDLEH GROUP

Insight

Information Solutions For You℠

Knowledge sharing for success

April 2006

Panning for Gold: The Search for Hidden Nuggets of Information

Consultant Roland Bernier may not wear a leather hat and vest or spend his life on the water, but every day that he goes to work at The MedLeh Group, he is panning for gold… searching diligently for those critical, hidden nuggets of information planted within the streams of data on clients' computer systems.

Obtaining the right information from thousands of files is not as easy as just putting in a few key words and pushing a button. It involves the careful analysis of client needs, utilizing specific techniques to expand and contract queries so that the search results are precise and concise, ready to use as supporting evidence in the specific legal cases for which they are required.

According to Bernier, there are two primary search models that are followed in electronic discovery. One is to handpick items to turn over to the other side, achieved by doing a narrow search with finite parameters. The other model is to look for items to exclude, such as confidential or sensitive documents or those covered by attorney-client confidentiality.

One of the reasons choosing the right search methodology is so important is that search parameters can greatly affect a client's budget and the way a case is approached by the legal counsel. All searches should be constructed with sufficient care that relevant data is not omitted from the results. However, the greater the number of documents to produce and review, the more costly a case can become, so a narrower search is typically better for a more cost conscious client. This is especially significant in electronic discovery, where there is a large volume of available data.

"The electronic discovery process opens the door up to a lot more data," says Carl Bruce of Fish & Richardson P.C. "Before, if you asked someone to get the relevant documents for a case, they would go to the ones that immediately came to mind or the ones that particular person thought were relevant and so you might only find 50% of the actual relevant documents. With electronic discovery, you can now find 80-90% of relevant documents."

Bruce explains that because search methodologies can impact the final litigation cost for clients, this is a frequent area of conflict between parties, where one side may be demanding broader search parameters in the electronic discovery process, thus running up the legal bill of the opposing party.

"It's going to be more difficult to control costs for clients when the new rules (on pre-discovery conferences) go into effect," Bruce says. "It will depend on what we're up against with the opposing counsel."

This is where The MedLeh Group can help. By serving our clients as consultants in the elec-

Panning for Gold continued on page 2

The Client Perspective

Speed and service is very important to us and so having The MedLeh Group in the same city as we are has been a real advantage. They are able to get the production material back to us immediately after it is sorted, which has helped us immensely when we were working with short deadlines. And because they have developed their own electronic discovery tool and are more flexible in their service offerings, they have been able to keep our production costs considerably under what the larger firms would have charged for the same jobs.

- Medleh Group Client

tronic discovery process, we can advise parties about the different options available to them which will broaden or narrow the results of a data search.

"Ideally, we would like clients to invite us to their pre-discovery conferences," Bernier says. "If they aren't willing to have us participate directly, the party doing the production should at least have an IT person there so they know what to ask us for."

The decision to utilize certain techniques is dependent upon the goal of the search; for example, whether parties are looking for concepts (requiring a broad search) or names (requiring a more narrow approach). How common the terms in the search are will also help determine the technique used. For instance, in searching for the name "Robert Smith", a search for a common name like "Smith" would produce too many irrelevant results and so a narrowing technique would need to be used in order to find the relevant files.

Some examples of techniques that help narrow a search include:

- Modifiers – utilizing adjacent names or adjectives such as first names with last names.
- Field searches – searching for something in a specific field, such a the date.

- Connectors – using "and" in a string, such as "Enron and Broadband" to define two words used together, or using "not" to rule out documents where two words are used together.
- Specifying the order in which words appear or how many words between them there must be.
- Weighted searches – scoring documents in terms of relevance to specific keywords.

Techniques that will broaden a search include:

- Fuzzy searches/pattern searches – looking for misspellings like an extra "n" in banana. The MedLeh Group's system enables us to "tune" the search from 1 to 10 depending on how far off the original spelling it will allow.
- Phonics searches – accepting alternative spellings of the same name such as "Smithe" and "Smythe" for Smith.
- Stemming – looking for different forms of the same word, such as verb conjugations and plurals; i.e., apply, applying, applies and applied.
- Concept searches – looking for multiple words that express the same concept, such as car or truck for automobile.
- Synonym and antonym searches.

In addition to helping clients define what search techniques need to be deployed in order to meet their objectives, The MedLeh Group takes several steps to ensure the process is as efficient as possible. First, the query is performed on the data before it is put into our EDD program, which saves time over querying it after it is already loaded. We also make sure the data query is performed prior to loading documents into iCONECT™, so that most of the "noise" is filtered out and clients only have to review relevant documents. Finally, the set of search criteria is saved as a file so that it can be used again in other cases.

For Further Insight

For more insight into electronic discovery methodology, visit the links below.

Law.com – Legal Technology Special Section on EDD
http://www.law.com/jsp/ltn/edd.jsp

http://www.discoveryresources.org/

http://www.ediscoverylaw.com/

http://www.lawyerlounge.com/ediscovery/index.php

http://www.arkfeld.com/

Pre-discovery Conferences Mandated in Amended Federal Rule

Electronic discovery is changing the face of the way cases are litigated because of the wealth of information that can be gleaned from computer systems and other electronic media. One result of this process change, however, is that disputes are frequently breaking out between parties over how data is to be managed and the way in which evidence is to be turned over. Disagreements between the parties over details in electronic discovery procedures can considerably lengthen pre-trial litigation, causing delays in the trial process.

Frustrated by these disputes, Federal courts have responded by amending Rule 26(f) of the Federal Rules of Civil Procedure to accommodate the new challenges posed by the use of information systems within the pre-trial discovery process. In the amendment, which is expected to take effect in December 2006, parties are mandated to meet and confer on the process to be followed. The rule expressly mentions that in these "pre-discovery conferences", the parties should try to agree on the form of production.

The goal of the amendment is to facilitate agreement on issues in order to hasten the pace of pre-trial litigation and reduce the expenditure of the court's energies on these matters. As such, it is not specific in its advice to disputants, allowing them to decide the issues that should be on the pre-discovery conference agenda.

Roland Bernier, one of The MedLeh Group's technical experts in electronic discovery, says some of the topics counsel might need to address in order to come to mutual agreement before electronic discovery begins include:

- Identification of potential sources of information and its associated custodians;
- The way in which the information will be turned over to opposing counsel (ie. forms of production such as log files, databases, etc.);
- Forms of media that might present challenges to production (such as obsolete backup tapes, etc.) and how the costs of pulling data from those media will be shared between the parties;
- How metadata will be handled;
- How embedded edits and similar properties will be handled;
- Methodologies for preservation, extraction, filtering and processing data for production;
- Identification of persons ultimately responsible and accountable for the preservation, extraction and processing of data for production;
- Procedures regarding inadvertent production and waiver of privilege, should confidential material accidentally be found and disclosed in the electronic discovery process;
- Identification of relevant subject matter and associated key words, including mechanisms for modifying the initial criteria; and
- Time frames for data production.

Bernier stresses that parties need to understand no generic list will adequately foresee all possible elements of different claims, and sometimes complex yet critical issues may arise out of these discussions. One such example is deciding what the nexus is between the producing party's key word list and its obligations to produce relevant documents. In other words, what is the obligation of the producing party if a document is relevant but does not

Pre-discovery continued on page 4

The goal of the amendment is to facilitate agreement on issues in order to hasten the pace of pre-trial litigation and reduce the expenditure of the court's energies on these matters.

Glossary

Search engine

A program designed to help find information stored on a computer system such as the World Wide Web, or a personal computer. The search engine allows one to ask for content meeting specific criteria (typically those containing a given word or phrase) and retrieves a list of references that match those criteria.

Query

A user's (or agent's) request for information, generally as a formal request to a database or search engine.

Fuzzy Logic

A type of logic that recognizes more than simple true and false values. With fuzzy logic, propositions can be represented with degrees of truthfulness and falsehood. For example, the statement, *today is sunny*, might be 100% true if there are no clouds, 80% true if there are a few clouds, 50% true if it's hazy and 0% true if it rains all day.

Fuzzy logic has proved to be particularly useful in expert system and other artificial intelligence applications. It is also used in some spell checkers to suggest a list of probable words to replace a misspelled one.

Relevance

In computer science, and particularly in search engines, relevance is a numerical score assigned to a search result, representing how well the result meets the information need of the user that issued the search query. In many cases, a result's relevance determines the order in which it is presented to the user.

Boolean Search

A query using the Boolean operators, AND, OR, and NOT, and parentheses to construct a complex condition from simpler criteria. A typical example is searching for combinations of keywords on a World Wide Web search engine. Examples: car or automobile, "New York" and not "New York state"

Proximity Search

A search option that looks for documents where the keywords are found within close proximity. It is used as an advanced search option in addition to searching for key words or searching for strings. Proximity search allows you to specify proximity relations between keywords of the search query. The proximity can be defined as number of words or number of characters. Sometimes also terms like NEAR, NOT NEAR, FOLLOWED BY, NOT FOLLOWED BY, SENTENCE or FAR can be used.

Wildcard Character

A special symbol that stands for one or more characters. Many operating systems and applications support wildcards for identifying files and directories. This enables you to select multiple files with a single specification. For example, in DOS and Windows, the asterisk (*) is a wild card that stands for any combination of letters. The file specification "m*" therefore, refers to all files that begin with *m*. Similarly, the specification "m*.doc" refers to all files that start with *m* and end with *.doc*. Many word processors also support wild cards for performing text searches.

Pre-discovery continued

contain any of the agreed upon search terms. Failure to discuss such an issue could significantly impact the outcome of a case or cause disruption and delays if not addressed until the trial was underway.

One of the ways in which The MedLeh Group can help its clients is to participate in pre-discovery conferences in order to facilitate communication between the parties on the technical details. Our experts can explain the specific methodologies utilized in the electronic discovery process, as well as providing time and cost estimates where parties need justification or are looking at cost sharing.

Attorneys seeking to better understand the parameters of the pre-discovery conference and issues beyond what are relevant to The MedLeh Group's services can refer to the Handbook of Complex Litigation.