



# Cyber Hunt for Evidence: Tales from the Trenches

It was an easy electronically stored information (ESI) investigation on the surface. There was no need to create an evidence image of computer hard drives; everything produced had been exported off of the company file servers onto DVDs for review. The only thing left to do was to mount the Microsoft Outlook e-mail backup files and search for evidence using 25 keywords provided by the in-house counsel.

Once the Outlook backup files were mounted, the search using the keywords began. Using only the first keyword, 500,000

individual e-mails appeared. That was too many to try to dissect to pinpoint evidence. Unfortunately, the legal team from the company had little working knowledge of computer forensics or data recovery. The keywords provided by the in-house counsel included the company name, a project name, common terms used in the company, and the names of individuals. Each e-mail had the company name in the signature block while an individual's name was repeated multiple times in original e-mails, forwarded e-mails, and company directories. The forwarding and replies on many e-mails increased the references of the one keyword by 200,000 thus expanding the e-mail population that required review to 700,000. In short, many keywords provided by counsel were useless in discovering relevant evidence for the investigation because the words put forth were too common.

Thus the keyword search terms provided by the legal team did not reduce the e-mail population to a manageable amount, nor were they designed to identify key e-mails that would assist the case. For example, an e-mail from one project manager stating "I'm not going to jail over this!" was not found using any keywords provided by counsel. An e-mail from a banker to a company executive asking why a million dollars had been taken out of a secure collateral account with a demand that it be "immediately replaced" also was not found. By modifying the original keyword list to words more likely to uncover

evidence pertinent to the case, those e-mails were identified. They became the start of the "clue trail" in which clue A led to clue B, which led to clue C, etc.

The foregoing example demonstrates the importance of consultation between client legal team leaders and forensic computer data experts at the outset of a claim to coordinate the ESI discovery effort. The goals and objectives should be clear; the facts counsel wishes to establish may not reside in the keywords provided but instead may be found elsewhere in the data. In the following cases, amendments to the keyword listing identified evidence that may have been missed with only the initial keyword search:

- In a marriage dissolution case, one party was certain \$1,000,000 had been hidden in a Swiss bank account by the other party. Therefore "Swiss Bank" was provided as one keyword by the client. It seemed reasonable to the client, but to the forensic examiner it was too limiting. "Suisse," "Banca," "Banque," "Svizzera," and "Schweiz" were quickly added to the list of keywords to ensure discovery of other possibilities. The examiner knew that French, Italian, and German are all common languages in Switzerland.
- The keywords provided to the computer forensic examiner would not have discovered that one party had searched the Internet for information on "Swift



**Brook T. Schaub**  
Contributing Author

Mr. Schaub is a retired police sergeant and seized computer evidence recovery specialist. He now serves as manager of computer forensics at Eide Bailly, in Bloomington, Minnesota.

MT999,” a type of verification message for proof of funds between banks (typically involving deposit amounts over \$1,000,000). The keywords provided also would not have discovered that the same party had searched for real estate in Costa Rica. Moreover, the romantic chat logs between that party and a “friend” containing discussions about bank accounts in the friend’s name, prepaid credit cards, and other currency-related transactions, may have been missed as well. The examiner’s knowledge of chat linguistics helped locate the evidence.

- In another case, deleted e-mail evidence was found by intentionally misspelling keywords in the course of the search process.

A report back from one’s adversary that the keywords provided revealed no responsive “documents” should not necessarily be taken at face value. There may be something sinister at work. Consider this example:

- During a civil case, one party’s counsel was advised that the keywords provided could not be found on the responder’s hard drive. After gaining access to the hard drive, the forensic investigation showed that the party’s IT manager had intentionally deleted and scrubbed files and folders from the disk two days before the court ordered that the computer be turned over for the requesting party’s examination. Needless to say, this information was extremely valuable from a spoliation of evidence standpoint.

In e-discovery forensics, examiners posit a distinction between “data recovery” and “data investigation.” The former involves the ministerial execution of database searches via keywords or document categories; the latter involves searches behind the keywords or documents through the use of metadata analysis and other investigations in search of the whole truth. Data investigation may require production of a bit by bit image of the opponent’s hard drive or other media devices. All hard drive or other media space, from the first “1” to the last “0” cell, should be imaged regardless of the number or size of the files present. The difference

between data investigation versus data recovery in the search for relevant evidence can be best explained through the following case examples:

- A business was in litigation with several employees who resigned abruptly. They were subject to noncompete clauses, yet the employees soon were employed by a competitor. Many keywords were used in searches of hard drives produced pursuant to a court order. Few of the keywords were of value in finding the truth. However, access to the hard drives allowed for examination of underlying data. That investigation showed that two days before the resignations, the employees downloaded customer lists, company strategic plans, pricing plans, market research, sales estimates, and 50 other company documents to a removable thumb drive. In addition, the deleted external web account e-mails present on a hard drive showed negotiations with the competitor by the employees prior to resigning, which included the employees’ knowledge of and concern for the noncompete clauses. The case settled quickly thereafter.
- In a company theft case, a picture file in an employee’s computer showed a couple enjoying a romantic beach vacation. Nothing appeared noteworthy, except forensic sleuthing found that one of the lovers was suspected of theft of company materials while the other was the comptroller in charge of the accounting software entries for those materials. The photograph explained why the accounting software “coincidentally” developed a destructive “virus” prior to an audit.
- In the course of the data recovery effort, a defendant company provided two gigabytes of data resulting in 1,000,000 separate data matches involving the 160 keywords provided. The matches included references in technical journals, news articles, company e-mail, website data, project plans, and more. The data match references were of little value in

the case. But using the two gigabytes of data produced, the computer forensic examiner employed data investigation techniques to identify the 12 files that were of value, without the use of keyword searches.

E-discovery forensic efforts add value to the search for the truth in civil litigation, now that ESI has supplanted paper-stored information in world communication and document preparation. In cases meriting forensic expert retention, early involvement

**A report back from one's adversary that the keywords provided revealed no responsive "documents" should not necessarily be taken at face value.**

of the investigator is important in addressing the search parameter decisions and articulating reasons supporting access to the adversary’s hard drives and other media repositories. The courts are increasingly ordering and encouraging counsel to form cooperative agreements for e-discovery processes. Attorneys need to be aware that these agreements must provide for flexibility, because access needs often evolve as the case work proceeds. Being able to zero in quickly on “smoking gun” data may eliminate further investigation, litigation, and expense to the client. Other values of forensic consultation are the examiner’s ability to provide guidance to counsel in drafting ESI-related discovery requests, and in articulating why the court should order the production of more computer hardware devices to ensure comprehensive gathering of discoverable information. Many courts lean toward limiting the scope of production given the creativity of counsel in offering arguments of undue burden, business disruption, and the always-present “fishing expedition” analogy. 