

LOS ANGELES

Daily Journal

FRIDAY,
MAY 5, 2006
Vol. 119. No. 87

— SINCE 1888 —

OFFICIAL NEWSPAPER OF THE LOS ANGELES SUPERIOR COURT AND UNITED STATES SOUTHERN DISTRICT COURT

Focus

Knowing Where, How to Look Will Cut E-Discovery Expense

By Bill Detamore
and Bob Dove

When a general counsel is presented with a bill for electronic discovery, a typical response is “How much?!” followed quickly by “What the #\$@!* for?”

Electronic discovery can be extremely large in scope, technically challenging and very expensive. Moreover, rapid changes in technology mean that costs incurred even a few months ago may be different from today.

In fact, if you are paying for the same process and costs as last year, you are probably not benefiting from advancements in technology.

Where is all this money on electronic discovery being spent?

The first phase of any electronic discovery project is collection of the data to be processed. Collection requires knowledge of the relevant players, where the data is stored, when the data is available and what type of content needs to be collected. Data collection is one of the most important phases of the process.

Costs for collection include the time for inside or outside resources to interview relevant witnesses (or “custodians”) and the IT staff that supports the infrastructure where relevant data is stored. For outside resources, this is typically an hourly charge and estimated based upon the number of witnesses and complexity of the request. If there are a lot of custodians, this could mean a lot of hours.

Collection of hard copy materials can be achieved usually by interviewing the creator or receiver of the information and

reviewing their relevant files. Collection of electronic data, however, also requires finding all of the places where that data may exist, including local machines, shared networks, mail systems and backups of all of those sources.

Casting too wide a net, however, can lead to collection of a lot of unresponsive data that must be separated from the truly relevant data, at additional cost. Using intelligent capture techniques and well-trained personnel can significantly reduce the amount of data to be processed and reviewed.

Collection may include identifying, collecting and processing backup tapes. Backup tape processing may be the single most expensive endeavor an organization faces in a discovery project. Backup tapes have the potential to store large volumes of data, and retrieving a tape does not mean that the tape is searchable, there may be additional costs associated with restoring it to a searchable format. Harvesting information from backup tapes may require enormous investment.

Sampling techniques can sometimes be used to reduce the amount of data to be processed from backup tapes. Responding parties may allow the requesting party to identify specific tapes, the parties may randomly sample tapes or the parties may sample tapes based upon a methodology that is intended to yield the broadest results in light of the organization’s tape recycling protocol.

Regardless of the method selected, it is advisable to reach agreement with opposing counsel regarding the sampling method early on and soon after the organization’s retention policies are known.

Sampling can be a significant cost reduction method for both parties.

Automated review processes such as culling or prefiltering are predefined, customized logic that identifies relevant documents prior to processing. Culling can include metadata, date and conceptual searching techniques. Applications that perform this process can search for specific date ranges or relevant search terms. Only documents that meet the criteria for potential relevance are collected for processing.

The use of such tools requires in-depth knowledge of the project. Prefiltering is effective only if you have a well-established list of search terms developed by staff with sound understanding of the allegations in the matter as well as a thorough understanding of the collection.

Altering search terms or re-processing with new search terms at a later date will add costs to the project. In addition to limiting the dataset to be collected, automated review applications can help limit the duration and expense of the human legal review.

Using the developed search term lists, documents can be preflagged for privilege, relevance or priority. This allows the documents most likely to be relevant to be reviewed first. Highlighting the search terms assists the attorneys and allows for the human review to be completed more efficiently and accurately.

Processing is converting the collected data into a format that can be reviewed by legal staff and produced to opposing counsel. The amount of data can be staggering. Forty or more boxes of hard copy information can be contained on a

single CD. A CD of common Microsoft Office files can generate hundreds of thousands of pages.

These numbers are much bigger than a typical hard copy job. The parties involved in litigation should meet as early as possible to determine how the data should be produced. Typical productions included native format, image with searchable text and metadata or a mixed format production.

Image and text generation along with metadata is a fairly common production practice. Every application creates different "metadata," which provides information about the file, such as how, when, and by whom a file was created. The creation of image files is an automated process of printing the data to image format instead of a printer. Printing hundreds of file formats in an automated and standardized process is a challenging endeavor, one that requires robust conversion software, and hardware and personnel that are proficient in many software applications, all of which add costs to your electronic discovery projects.

A cost that is often overlooked in processing is the quality control effort. In the world of scanning hard copy documents it was standard practice for service bureaus to review every page that was imaged. Because of the sheer volumes of data being processed today, a page level review in electronic data processing is not reasonable.

However, quality control still exists and it consumes a large amount of time. At a minimum, a random sampling quality control of standard file formats should take place.

An additional complicating factor is that many documents, such as databases or large-format spreadsheets were never intended to be printed. A process for handling these documents must be identified and implemented before processing begins. If image generation is required, expect many billable hours of manual formatting.

Processing for native file review is less labor intensive than image generation and typically quicker to deliver data for the review team. The conversion team produces the native file along with the metadata and searchable text.

This step is relatively quick and inexpensive. However, the review of native files may be more expensive than reviewing images. Reviewing the files with the application that created them requires a copy of the software for every possible file format to be reviewed. In addition the review staff needs to be proficient in multiple software programs.

Alternately, native review can be accomplished with a universal viewer, which is an application with the ability to render an image or html view of a document without having the native application loaded on the machine. This eliminates the need for the native application and proficiency in multiple applications.

Reviewing images can be more efficient than native file review. The image is a standard format that is easily opened with many image viewers. Most image viewers have similar controls, once an operator learns how to navigate the images it is the same process for every document regardless of the original file format.

While many e-discovery costs are unavoidable, the following is a condensed list of actions that can significantly reduce them:

- Agree upon a production format as early as possible.
- Develop and implement a collection strategy that will collect only relevant material and limit disruptions to employees' normal activities.
- Generate images only if they are required for production or the review process. If generating images, define exception handling procedures for file formats that are not intended for printing.
- Further reduce the data set prior to legal review by taking advantage of culling, prefiltering and possibly sampling techniques.
- Implement or use a review system that is robust enough to handle multiple review formats while managing the review workflow.

Electronic projects typically are much more voluminous than traditional hard copy projects. In addition they can be technically challenging. Planning for electronic discovery and implementing some of the strategies discussed can assist in keeping

costs under control.

Other practical tips for reducing e-discovery costs include:

- Defining the relevant time frame. Don't spend time processing and reviewing documents that are not relevant. Identify a time frame and filter out any documents that do not meet the requirements.

- Identifying the key personnel; whose data should I be reviewing? Learn where key personnel store their data and with whom they communicate. This will assist in collecting only the relevant materials. Identifying relevant search phrases for relevance, prioritization and privilege reviews. Using automated review software, you can reduce the amount of data that needs to be reviewed by a human review team. Priorities and privilege status can be applied to the documents that need to be reviewed by the legal team. This identification will make the review faster and more accurate.

- Encouraging sampling. If you have to restore backup tapes, encourage sampling. Favorable results from a sampling test can limit your obligation to search and restore additional back up tapes.

- Defining the production format as early as possible

- Knowing your infrastructure. Learn and understand your network infrastructure. You need to understand what storage locations are available and who is using them. The majority of electronic material is stored and or transported via e-mail. Learn the e-mail system and policies for e-mail storage.

- Knowing your document retention policy (or get one!).

- Knowing the back up rotation — for file servers, e-mail servers, duration, rotation, tape format, archive software, etc.

- Knowing your IT staff. The IT staff will be a valuable resource when it comes to identifying and retrieving electronic material.

Bill Detamore is CFO of IE Discovery Inc., an Austin, Texas-based discovery management company with additional offices in Cypress. E-mail bdetamore@iediscovery.com. Bob Dove is the e-discovery consultant at the firm. E-mail bdove@iediscovery.com.