

The e-Discovery Chase: Strategies to Reduce Electronically Stored Information in Discovery Requests

All of our tools and toys that generate electronically stored information can frustrate law firms with high e-Discovery costs. Lawyers who choose to have their discovery productions printed can have nearly 10,000 times more paper than 10 years ago.¹

It is difficult to visualize how “big” ESI can be. WIRED magazine recently gave examples of data size most people can comprehend²:

- 1 Terabyte: a Hard Drive with 260, 000 songs
- 20 Terabytes: All the photos uploaded to Facebook each month
- 120 Terabytes: All the data from the Hubble Space Telescope
- 530 Terabytes: All the videos on Youtube
- 1 Petabyte: Data processed by Google’s servers every 72 minutes
- 12 Exabytes: All human produced information³

The data explosion can drive up e-Discovery costs for collection, processing and review to rival that of fielding an America’s Cup team. And like the America’s Cup, if only billionaires can afford access to Federal Courts, then justice is not being served.

Danger of Costs from Overly Broad Requests for Production

Given the sheer volume of ESI, discovery requests need to be focused to avoid opening a Pandora’s Box of data. For example, a third party request for production demanded the following:

The content of any and all electronic files, e-mail messages (with attachments), Instant Message communications and/or other communication created any time between August 20, 2001 to July 20, 2007 and maintained by Yahoo! related to account holder Jacqueline Hone’s subscription with Yahoo!, Yahoo! mail and/or Yahoo! Messenger.⁴

Even if the court had allowed production of this electronically stored information, how is it a victory to review SIX YEARS worth of email and instant messages? The cost to collect the information, process it and review time would make the venture questionable.

Lawyers need to employ strategies to reduce the volume of ESI to control costs. These strategies including requesting what you need, utilizing trusted vendors for targeted collection, and using technology for pre-processing analysis, effective processing and intelligent review.

The Discovery Request

There is no discovery production to review without first a discovery request. Lawyers should avoid discovery requests such as, “All email to or from the Plaintiff from 1999 to 2008.” Such requests are rarely granted when challenged, because they are overbroad, unduly burdensome, and often amount to a mass fishing expedition.⁵ Moreover, getting 800,000 email messages from one individual would be a Pyrrhic victory considering review costs.

Lawyers should consider timeframes, form of production, specific individuals and other factors to create a narrowly tailored request. This may reduce motion practice and help control production and review costs.

¹ John Bringardner, WIRED, “Winning the Lawsuit,” page 112, July 2008

² “The Petabyte Age,” WIRED, July 2008, pp 106-107

³ <http://en.wikipedia.org/wiki/Exabyte>

⁴ *Hone v. Presidente U.S.A., Inc.*, 2008 U.S. Dist. LEXIS 55722 (N.D. Cal. July 21, 2008)

⁵ See, *Hone v. Presidente U.S.A., Inc.*, 2008 U.S. Dist. LEXIS 55722 (N.D. Cal. July 21, 2008), *Quinby v WESTLB AG* 2006 WL 59521, 1 (S.D.N.Y) (Jan. 11, 2006), and *Thompson v Jiffy Lube International, Inc*, 2006 WL 1174040, 3 (May 1, 2006).

A party responding to such an overly broad request should object and force the demanding party to focus their requests for “all email” or “all information related to the website.” The petabyte explosion of ESI has not eliminated the discovery requirement that requests be narrowly tailored and reasonably calculated.

Targeted Collection

Information defensibly collected off hard drives should be focused on relevant or responsive electronically stored information. If representing a contractor in a construction defect case, targeting collection to the housing project at issue will collect less information to review than a mass copying of hard drives. However, care should be given not to under collect ESI, creating a risk for re-collection if something is missed. Moreover, this should not undercut any preservation duties. One might have to mirror image a hard drive to preserve all the data if the facts call for it, I'd say lose this comma but what is collected for review should be focused.

Pre-Processing Analysis

Software tools have been developed for “non-linear” review for “pre-processing.”⁶ What this translates into is looking at email messages before they are processed⁷ or prepared for a review system⁸.

“Pre-Processing” allows a look at email strings and a determination of what needs to be processed for linear review in a product like CT Summation iBlaze or Lexis Nexis Concordance. This stage can eliminate email messages that are spam or newsletters based on their domains, such as “ebay.com” or “nytimes.com” to reduce what is ultimately loaded into a review platform.

Effective Processing

Processing of electronically stored information is the extraction of metadata and full text in a format readable in a review platform. The number of native files ESI processing software can process are in excess of 500 to 700 native file formats and growing.

Processing can be “brute force” and literally just included everything that was collected. However, just because preservation duties may require the mirror image of a hard drive, that does not require an entire hard drive be processed for production. Processing can be focused with key words, such as names of parties or witnesses, date ranges and other criteria to narrow the information for review. Such “smart processing” can reduce costs by shortening review time by narrowing the production to a focused production.

Making the Most of Review: Not just a box of paper

Reviewing ESI is not the same war room experience as digging through boxes of paper. Conference rooms do not need to be commandeered for months with contract attorneys plowing through fields of paper for the smoking gun document. Technology can help reduce such backbreaking work.

Law firms can avoid the haphazard approach to review with assigning reviewers “review sets.” A “review set” is saved database search, that can be a DOCID, date range and keyword(s). The litigation team generally knows some of the basic terms and dates relevant to the lawsuit, and searching based on these terms for review helps focus review. Reviews can also be set for specific individuals, document types or almost any term appearing in the database.

Review tools allow for issue coding (and also vary by product). Coding for causes of action, specifically by elements of causes of action, allows for thoughtful review and case preparation. A

⁶ Companies such as [Clearwell](#) and [Metalincs](#) both make products for this purpose.

⁷ Processing products include [IPRO's eCapture](#), [Needle Finder](#), [LAW Pre-Discovery from Lexis](#) or [CT Summation Discovery Cracker](#), to name a few.

⁸ Review platforms as hosted solutions such as [Concordance FYI](#), CT Summation CaseVantage or desktop solutions such as [CT Summation iBlaze](#) or [Concordance](#), among many other fine products.

reviewer can also consider admissibility issues during the coding stage, identifying any problems before trial.

These are just a few strategies for review. There are many others. The main point is not to treat fully searchable electronically stored information as a box paper. Searching and organizing based on key words can help focus review into what is relevant.

Don't Drown in Petabytes: Search to Reduce ESI

All of our cool iPhones, laptops and Wii's are engines of electronically stored information. Technology has created an excessive amount of ESI in litigation, but technology can solve the problem in reducing ESI so law firms can focus on practicing law, not drowning in data.