

Minimizing the E-Discovery Spend: Leveraging Technology with Better Lawyering

By Anne Kershaw

For the last several years, in-house counsel have been bombarded with a seemingly endless assault of ads, articles, seminars, webinars, emails, blogs, podcasts and tweets – all proclaiming to have effective technical solutions for quelling the rising costs of dealing with the discovery of electronically stored information (ESI). To be sure, certain technologies *can* have a dramatic positive impact on reducing the cost of handling ESI, often cutting the cost to review the data, the most expensive part of ESI discovery, by over 90%.

However, often overlooked in the hyperbole about technology is the fact that effective lawyering can also dramatically lower the costs of electronic discovery, principally by better targeting the scope of ESI that is collected in the first place. Proactively engaging the court when your adversary refuses to discuss a sensible and targeted preservation and collection plan is also essential – judges have become much more sensitized to the costs and problems associated with ESI and we lawyers need to do a better job of engaging their assistance.

In other words, don't just talk about proportionality, walk the walk by learning and truly understanding the most informative client data sources and educating your adversary and the court, if necessary. This, coupled with appropriate objections regarding collection from redundant sources, letters such as the exemplar set forth below, and a custodian-centric collection process that includes robust custodian interviews, can go a long way in reducing the amount of ESI you need to collect. Indeed, the rise in popularity of the topic of proportionality reflects acceptance by courts and commentators that appropriate preservation and collection efforts are those that are informed and targeted, acknowledging that the concept of "reasonableness" includes an understanding that even if mistakes in good faith are made, the true question is whether the complaining party can show that something that mattered to the case is missing. See



"[O]ften overlooked in the hyperbole about technology is the fact that effective lawyering can also dramatically lower the costs of electronic discovery, principally by better targeting the scope of ESI that is collected in the first place."

Orbit One Communications, Inc. v. Numerex Corp, 2010 WL 4615547 (SD NY, October 26, 2010).

Discussed first below are the steps that lawyers can take to limit what is gathered, followed with a review of the types of technology or processes that lower costs and improve the quality of document review.

Lawyering

Here are some simple, practical steps that can be taken to appropriately target the volume of ESI to be processed and reviewed:

1. Issue a general litigation hold notice within the organization and send a letter to the adversary or government advising that, among other things, you will be conducting a custodian-centric preservation and collection process and that you will seek to avoid the collection and production of redundant data. See exemplar letter below.
2. Understand what systems and enterprise applications the company uses in the business areas affected by the case, how information is entered and stored in these applications, and who uses them for what reasons. Ask IT to identify routine deletion processes and suspend any that might affect relevant information.
3. Identify and document key employees and documents - learn the terms those people used to discuss the matters at issue, the people and organizations with whom they communicated on those issues, and the form those communication took. Note: These initial "preservation scope" interviews will be followed up with more detailed interviews at the time of collection.
4. Using the information from steps 2 and 3 above, negotiate a sensible discovery plan with the adversary that includes an in-depth analysis of the

Anne Kershaw, “Minimizing the E-Discovery Spend: Leveraging Technology with Better Lawyering”

value of data reports from company systems and applications, together with a tiered process that excludes, for example, electronic files from the initial collection because they are largely redundant of email attachments. The discussion should include reference to the technologies discussed below and agreement should be reached on the form of production (with consideration of the lower cost of on-line productions). An agreement on privilege waiver and logging should also be reached (to be so-ordered per FRE 502). If the adversary will not negotiate, present a proposed plan to court. Note: in some cases this negotiation may be greatly assisted by handing over at the time of the discussion the data and documents that you have already located and determined to be relevant.

- For purposes of targeting the collection, engage a robust and documented interview process of the core group of key players identified in step 3, gathering their relevant paper and electronic data and analyzing it, looking for other names, email addresses, positions, organizations, etc. with whom the key players communicated. Use this information to identify additional past and current employees to add to the group. Data collections pertaining to past employees or legacy data will involve separate processes (to be discussed in subsequent articles).
- Repeat the process until it appears that the records identified to date represent a reasonably complete collection of relevant records along with a useful compendium of concepts and terms used.
- Document each step and preserve all the records gathered. Use the information from steps 5 and 6 to issue an informed and targeted litigation hold notice (to replace the initial broad notice that was issued when the matter came in).

The simple expedient of gathering relevant and accessible records and telling the other side about it defuses expensive Monday-morning quarterback attacks by opposing counsel and courts later on in the case. Handling ESI without this transparency and without the willingness to go to the mat over what is reasonable or proportionate at the outset in any given situation can result in extensive over-gathering and processing in anticipation of being challenged at some later date on the defensibility of the production. Trial counsel simply have to go hands-on with the data early

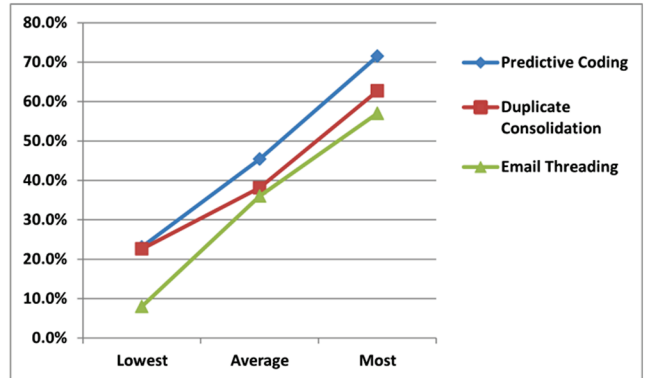
in the case and charter an aggressive course of not going beyond what is reasonable in the light of all the circumstances.

Technology

Having conducted a reasonable but focused collection of records, now is the time to gain the benefits of technology to leverage the effort required to review those records prior to production. The “Judges’ Guide to Cost-Effective E-Discovery,” published by the eDiscovery Institute, a 501(c)(3) nonprofit research organization, summarized its earlier surveys and studies and recommends the following technologies and processes¹:

DeNISTing – the National Institute of Standards and Technologies publishes a list of all known files associated with commercial software along with so-called “hash” values that act as electronic fingerprints to assure that the actual software files are located, not just files that appear to be software files. Unless the litigation concerns some of that commercial software there is no reason to gather software-related files. With some e-discovery fees based on number of gigabytes or number of files collected, programmatically eliminating software and system files can cause a significant reduction in costs.

Average Responses to Surveys by EDI on Lowest, Average and Most Savings Observed Through Use of Across-Custodian Duplicate Consolidation, Email Threading and Predictive Coding



Duplicate Consolidation. Technology now allows us to eliminate duplicates without losing any information as to the sources of those duplicates. This means that only one copy is necessary and having reviewers examine only one copy of each email or document can save, on

¹ The Judges’ Guide is available in electronic form at no charge at www.eDiscoveryInstitute.org/JudgesGuide.

Anne Kershaw, “Minimizing the E-Discovery Spend: Leveraging Technology with Better Lawyering”

average, 38% of the cost of review. Some systems will only identify duplicates within but not across the records of multiple custodians, causing redundant reviews, e.g. if an email goes to four people it may be reviewed in the records of the sender plus reviewed when each of the recipients produces it. This double billing exposes the corporation’s records to the risk of inconsistent review decisions as different reviewers can make different decisions on different copies of the same records. It also needlessly permits additional people to read confidential company records because of all the unnecessary duplicative review. For a more detailed examination of the ethical implications of failing to consolidate duplicates, see “Ethics and Ediscovery Review,” by Patrick Oot, Joe Howie and Anne Kershaw, *ACC Docket*, Jan/Feb 2010.

Email Threading. An email thread is formed when a recipient of an email replies to or forwards an earlier email that he or she received. Email threads can continue for as long as the senders and recipients continue the conversation with additional replies and forwards. As ought to be obvious, reviewers who are presented with all the emails in a thread at the same time will be able to comprehend the entire thread much easier and more quickly, and in fact the average savings from using threading technology is 36% beyond the savings achieved by duplicate consolidation. As with duplicate consolidation, not only are costs lower, but the quality of the review is higher as the reviewers can make relevance and privilege determinations based on all of the content and recipients in the whole thread.

Predictive Coding. Predictive coding or document categorization involves reviewing a subset of the entire record population and then propagating decisions made on that subset to the entire population. The benefits of this type of technology are of two types: (1) some records may be included in the production or excluded from it without having to have someone read them, and (2) records can be ranked by likely relevance with the most relevant assigned to the most knowledgeable reviewers and the least relevant assigned to the most junior members of the review team. While there are many ways to do document categorization, the average savings of such approach is 45% beyond that achieved by duplicate consolidation and email threading.

While some lawyers are understandably concerned about the quality of such a technology-assisted review, studies have shown that the long-used manual pre-

production review of records for relevance is itself highly suspect. As detailed in the Judges’ Guide, studies have shown that a record selected as responsive by one team of reviewers will also be selected as responsive by a second team of manual reviewers in only 48-62% of the time -- about the same odds as a coin toss. At least with predictive review, if the same process is applied consistently to the same records, using the audit trail created by the first review, consistent decisions will be made.

Conclusion

As shown above, to realize meaningful reductions in the costs of e-discovery, corporations need to have counsel willing to use their trial lawyer skills not just to win the case but also win in obtaining agreement on, or a court order for, an informed, sensible and targeted preservation and discovery plan. To do this these lawyers will need to learn the corporation’s processes for creating and maintaining its business information, the people involved in those processes, and how this information interplays with the allegations made in the case. Establishing and holding the line on a reasonable preservation and collection plan and process is essential to securing proportionality and holding down costs. Once records are gathered, various technologies such as deNISTing, duplicate consolidation, email threading and document categorization can significantly lower the cost and improve the quality of the pre-production review.



Anne Kershaw is founder of **A Kershaw // Attorneys & Consultants** a litigation management consulting company providing innovative and impartial analysis and recommendations for the management of all aspects of litigation. Anne is also the co-founder and president of the **Electronic Discovery Institute**, a 501(c)(3) nonprofit research organization that conducts studies and industry surveys of technologies and processes for processing ESI. Anne used the experience and insight from her work with the Institute in teaching workshops on e-discovery at the **Federal Judicial Center’s 2010 Conferences for Magistrate Judges**.

Anne Kershaw, “Minimizing the E-Discovery Spend: Leveraging Technology with Better Lawyering”

Anne is a member of the **Sedona Conference®**, a Faculty Member of **Columbia University's Executive Master of Science in Technology Management**, is on the Advisory Board and faculty for the **Georgetown E-Discovery Training Academy**, and is an **Advisory Board Member of the**

Georgetown University Law Center's Advanced E-Discovery Institute. She also co-chairs the **ABA's E-Discovery and Litigation Technology Committee for the Corporate Counsel Litigation Section Committee**.

You may contact Anne at Anne.Kershaw@AKershaw.com.

Sidebar 1: *Letter to Opposing Counsel Describing Reasonable Efforts to Preserve and Produce relevant information.*

Re: Records Preservation and Production Efforts in X v. Y litigation.

Dear Ms. Smith:

Pleased be advise that we are in receipt of your [Complaint] in the above-referenced matter, service of which on Defendant _____ is hereby acknowledged.

Please be assured that we will take all necessary steps to preserve information that we understand, based on the information you have provided, is relevant to your claims and our defenses. Our usual process for satisfying our preservation obligations in civil actions such as this is through the issuance of a notice to preserve to all employees believed to have knowledge regarding the location of relevant information. As additional information is obtained regarding the location of relevant data, the scope of our initial preservation notice may be broadened or narrowed.

In addition, we will endeavor not to preserve redundant or cumulative information and accordingly, we will not be using our backup disaster recovery system, legacy backup tapes, or backups made by employees outside of the company's backup system, to preserve data for this litigation. Similarly, absent special circumstances, we will not be imaging employee desktop or laptop hard drives or drives in network servers. Information collected for other matters will not be preserved for this litigation unless it involves the same or similar claims or defenses. Voice mail will be preserved as it is received, if relevant, on a going-forward basis. Any user created databases or email archives will be identified through employee interviews. Finally, if there are particular log files associated with operating systems that you feel must be preserved, please notify us immediately.

Please notify us immediately if any of the forgoing is unclear or if you would like to discuss alternatives for preserving relevant data.

Sincerely,

Anne Kershaw, “Minimizing the E-Discovery Spend: Leveraging Technology with Better Lawyering”

Sidebar 2: *Letter to Outside Counsel Inquiring About the use of Cost-Effective E-Discovery Processes.*

Re: Cost-Effective E-Discovery Processes

Dear XXX

I would like to become more familiar with the processes and technologies that you are using to reduce the cost of processing electronically stored information. My research indicates that many of these cost-reducing processes not only reduce costs but reduce the number of people who have to be given access to our records and minimizes the risks of inconsistent review decisions.

Please advise me as to your use of the following technologies or processes along with some metrics on the savings that they are achieving:

- **deNISTing** – are you using procedure to not copy software-related files, focusing instead on user-generated data?
- **Duplicate Consolidation** – in particular I would like to know if you are consolidating duplicates across custodians or just within individual custodians.
- **Email Threading** – are reviewers presented with all the emails in a thread at the same time?
- **Predictive Coding or Document Categorization** – are processes in place to permit the production of electronic records without having to visually examine every one or to have the most likely relevant records reviewed by the most senior people and the least likely by the most junior?

Information on these technologies and processes is available in the Judges’ Guide to Cost-Effective E-Discovery, published by the eDiscovery Institute and available in electronic form at no cost at www.eDiscoveryInstitute.org/JudgesGuide.

Sincerely,

VP and GC