



**Of Swamps and Alligators:
Avoiding the Bite of a Poorly Executed Response for ESI**

May 10, 2017

Of Swamps and Alligators: Avoiding the Bite of a Poorly Executed Response for ESI

Your client, a pharma company, just received notification they are being investigated for off-label marketing for one of their drugs. It appears a former company salesman approached the regulator as a purported whistleblower. The Subpoena lists several categories of information. The waters could get deeper - potential criminal investigation and Medicare fraud. And the first of product liability claims from patients is filtering in.

In-house counsel is litigation savvy but not well versed in eDiscovery management and tells you they want to save costs and have company IT staff conduct targeted collections.

Electronically Stored Information that Satisfies Proportionality

The best chances to avoid loss of limb are to satisfy the regulator by a complete production to explain the drug promotional activities at issue, often done by a maverick sales rep or a District Sales Manager wanting to increase territory revenue. Best case scenario for the investigation - the regulator makes an informed assessment on the scope and level (or lack) of management knowledge/responsibility and moves on with minimal fines. However, the product liability claims add a wrinkle. Now the challenge is to maximize efficiency of data preservation and data collection efforts while avoiding duplicative collection, hosting, and review efforts to the extent that's possible.

Transparency - of how you helped the company make defensible productions leveraging technology to address all areas of regulatory inquiry - will go far in securing a favorable outcome for your client. You're sensitive to client cost concerns and seek to ensure discovery efforts to yield a relevant set of ESI that also satisfies proportionality principles.

Alligators anticipate an opportunity in this swamp of legal technology and eDiscovery. Here are potential electronically stored information missteps and how avoid an attack.

Potential ESI Bite #1: Data Preservation Missteps

As a starting point, organizations are permitted to have reasonable record retention policies. However, data preservation obligations during the pendency of litigation/investigation for custodians and to relevant data sources require an investment of effort from the get go. A few key measures can be

employed to show good faith cooperation in investigations and avoid costly discovery motion practice on proportionality in the civil litigation context.

Given the potential for this matter to morph into more, you can help your client by having a sound data preservation plan, keeping these thoughts at the forefront:

- Legal holds (and custodian reminders) are key! Sanctions and adverse inferences are routine for oversight or tardiness of hold issuance.
- Involve, and involve early on, the company's IT and Record departments in the data preservation process to suspend routine document destruction and suspension of automatic email deletion.
- Custodian interviews enhance not only an understanding of custodian roles and issue familiarity, but also strong keyword development for identifying document sets potentially responsive to requests for production. Reiterating preservation duties at this point can go a long way to minimize discovery about discovery.
- Assess if it makes sense to keep preservation copies of strategic back-up tapes as a prophylactic measure. The mantra for good ESI practice goes: "Preserve broadly, collect narrowly."

Potential ESI Bite #2: Art and Science of Data Collection Targeting

Voluminous ESI (electronically stored information) is a fact of life and here to stay. Legal practitioners, however, have a toolbox at their disposal to help clients defensibly narrow the volume of data collection.

Some of the more important data collection targeting tools are:

- Stipulate to agreed-upon facts so discovery only occurs for disputed facts. *This is an often overlooked, low-tech solution.*
- Spend time to focus on what's going to be relevant, what the claims are, and what the defenses are.
 - Based on that analysis, can you exclude certain data types? Audio or video files? Others?
 - Can certain data sources such as disaster recovery backup tapes and archive data, or legacy systems be seen as sources of cumulative, non-unique ESI, and therefore be excluded? If not, retain a vendor that can produce operational metrics related to the costs of ESI collection for your proportionality-based defense.
- Explore if your client has enterprise archiving or journaling of data. Clients can then leverage their own technologies, thus streamlining litigation costs and defensibly excluding local stores of duplicative data.
- Advanced technologies, such as early case assessment by the client to quickly search their own databases, are gaining popularity. This would allow the client to develop defensible data culling processes that identify sufficient (but not over-inclusive) ESI collections.

- Your list of custodians will invariably be challenged with the argument that additional individuals have additional information. You can counter in a variety of ways with likely mixed success given the circumstances of the case and how courts have applied proportionality principles to similar motions.
 - The decisions turn on an analysis of various factors:
 - The personnel at the company involved with the issues of the case (the larger the number the more likely courts are to permit additional custodians considering these to be proportionate to the size and scale of the action).
 - The more detailed the description by the requesting party to justify their request for additional custodians, the more likely a court is to grant it.
 - Courts are aware of electronic de-duplication technologies and have cited to these as ways for defendants to limit review and production of duplicative documents, reducing some of the burden on these parties of producing information from additional custodians.
 - Courts also look at the size of the companies and their resources.
 - Watch out for sensitive data sources such as ESI protected by foreign privacy directives, data containing PHI or PII. The costs associated with compliance may prove onerous.
 - Date range filters – one size may not fit all! Explore if narrower time frames apply for certain claims to narrow whenever possible.

Potential ESI Bite #3: Too Much Review Data

From collections to review. One decision by Judge Peck is instructive of several practice pointers when it comes to narrowing down the collection set to arrive at the review set. Judge Peck refused to force a party to use technology-assisted review (TAR) rather than keyword searching in *Hyles v. New York City* (S.D.N.Y., 2016). He stated that under Principle 6 of the Sedona Principles for addressing document production “responding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information.” He added “some vendor pricing models charge more for TAR than for keywords. Usually any such extra cost is more than offset by cost savings in review time.”

Judge Peck’s dicta speaks volumes.

- First, find an ediscovery vendor that offers competitive pricing for TAR on a review platform that works for your case team.
 - Side note – what exactly is TAR? Ensure you understand these terms and what practical applications work for your data set. Technology Assisted Review or TAR extends to many helpful technologies that find patterns in data. Leveraging the patterns helps us review and more swiftly understand and review the deluge of data we’re awash in.
 - **Structured Analytics** broadly translates into email threading and duplicate identification; “structured” because data grouping and patterns are based on the data text or the metadata structure. No algorithm, no judgment, no machine learning – just structure!
 - **Conceptual Analytics** utilizes an algorithm to conceptually group data either with or without human input across the data of interest and using adjustable

similarity settings. “Conceptual” because based on concepts regardless of text and structure; algorithm, predictive (not descriptive), machine learning!

- Next, must you always use TAR? Which TAR techniques can always benefit a case? Which are elective? And must you track or create supporting documentation? As Judge Peck notes, you (in consultation with client/ custodians/ consultants) are best suited to determine searching methodology based on your understanding of the data.
- Keyword usage as a methodology to narrow the potential relevant set of documents for review and production is widely accepted. Experience and research, though, have shown keyword searching is imperfect and has the “Go Fish” problem (where it becomes a guessing game).
 - Developing effective search terms should be an iterative process that starts with input from knowledgeable custodians and individuals and a review of representative documents.
 - Next is testing and refining the search terms by using available technology to its greatest advantage. Such an iterative process may serve as a significant advantage if searches are challenged. You may need to offensively or defensively revise keywords (either to contract or expand terms) over time as the matter progresses and you learn more about the case, the documents and the terms used in the ESI.
 - The “Go Fish” problem can be lessened if you understand the limitation of search terms and broaden to concepts. Search term expansion can occur with the use of TAR by including conceptual synonyms. Predictive coding and other analytics-based functionalities all can help identify responsive files.
 - Getting party agreement early on in the “meet and confer” on the iterative process in developing search terms is critical to managing the risks and costs associated with a large document review.
- Sampling is often used to great effect to test terms and also to provide to opposing parties to satisfy them of the sufficiency of terms.
- Finally, documentation and metrics on keyword development provide a strategic advantage. Such reports typically would include search term results, unique hit counts, and reason for ultimate selection of terms. Many cases turn on such reports (or the absence of such reports!).

Potential ESI Bite #4: Too Many Hours of Document Review; Production Data Dump

Chances are you will be faced with the large expense line items of document review and document production.

The most dramatic and positive effect of the use of TAR can be seen in both improving and speeding up document review. It does this in several ways:

- Prioritization of review using conceptual analytics and predictive coding accelerates the time in which you get your hands on key documents. Clustering documents can occur without any reviewer input to group conceptually similar documents allowing your team to get to key documents faster and dispense with “junk” documents.
- Predictive coding in lieu of a linear contract attorney based model dramatically cuts down on costs is gaining wider acceptance in civil litigation and government investigations.

- Using TAR also cuts down on the inconsistencies seen in manual review which some studies on the extreme low end place at 84% (with agreement rates at 16%).
- If you choose a manual review, review software and vendors can assist with ensuring that you reap the benefits of grouping documents by email threads or near duplicates for swifter and more consistent coding of documents.

Production costs can similarly be corralled.

- Electronic repositories are a great way to exchange data. They also give you greater control in executing on privilege clawbacks.
- Carefully craft the production specifications. For example, you neither want to produce Excels in voluminous images nor do you wish to receive such imaged spreadsheets.
- If PowerPoints are key for your case be sure to request and produce in native format.
- You do not have to bow down before arguments that such file types have to be produced in image format either because they cannot be endorsed or because they cannot be redacted. The technology is here for that!

If you're asked to gather and produce additional documents don't panic. There is help for that, too.

- You can argue that the cost of the production, when considered against the amount in controversy within the case, renders production "unduly burdensome." To be successful, be prepared to show cost breakdowns for processing, searching, reviewing and producing the requested ESI, along with other cost estimates for alternative search parameters based on the plaintiff's discovery requests.
- You would need to tie costs to what is at stake in this litigation to argue that the additional demands are unduly burdensome. Showing the amount of ESI you've already produced combined with the high costs of producing the additional ESI should be helpful.
 - You can request the other side fund the costs for additional production but then be sure to include time for attorney supervision and QC review, and not just initial document review costs.

Discovery that is redundant, disproportionate, or results in a fishing expedition is likely to be seen as overly burdensome. Analysis at the critical points and utilizing technology for forecasting costs can be invaluable to contain costs. You have a duty of technical competence to your clients. Given the myriad issues this duty can be satisfied in part by consultation with eDiscovery consultants who can guide and support your decision making.

To summarize, your ESI responses should be framed around the guiding principles of proportionality and, if thoughtful inquiries and foundations are applied, you should end with a drained swamp, a happy client and hopefully all your limbs.

Mimi Singh is the Associate General Counsel and Director of eDiscovery Consulting for Evolver. She provides full-service e-discovery lifecycle consulting and project management relating to identifying, preserving, processing, reviewing, and producing in an efficient and cost-effective manner. She also consults with clients on a more automated method for reviewing, redacting, and producing native excel files with Evolver's XLeator Excel redaction software. Evolver provides free Relativity structured analytics - email threading, language identification, textual near duplicate identification – to clients to assist them with workflow efficiencies.