Portfolio Media. Inc. | 111 West 19th Street, 5th floor | New York, NY 10011 | www.law360.com
Phone: +1 646 783 7100 | Fax: +1 646 783 7161 | customerservice@law360.com

Time To Ditch Traditional Methods In Merger Probes

(June 6, 2014, 10:42 AM EDT)

Jonathan L. Lewis

Organizations facing the prospect of a merger investigation — as well as those organizations' attorneys — should understand and consider alternatives to traditional document-by-document review. But don't take our word for it.

In a recent publication titled "Technology-Assisted Review and Other Discovery Initiatives at the Antitrust Division," U.S. Department of Justice Antitrust Division's senior litigation counsel for electronic discovery, Tracy Greer, highlights an ongoing division initiative, begun in March 2012, to explore and encourage alternative methods of reviewing documents and preparing responses to second requests.

Parties to merger investigations have long struggled with the burden, cost and time necessary to respond to investigators' requests — and those struggles have only increased with the proliferation of electronically stored information.

Greer's paper focuses on processes and tools to reduce the costs of separating relevant from irrelevant data and reducing the "size of document productions received by the division, without undermining the ability of the division to conduct an appropriately thorough investigation."

While the publication's primary focus was on expanding the use of technology-assisted document review (TAR) within the investigative context, there are other improvements that organizations may utilize as well, provided they are disclosed to and discussed with the division.

Since the advent of specialized e-discovery practices, experienced outside and in-house counsel have gotten much better at identifying likely sources of information, applying date and location restrictions, employing deduplication and culling measures, and even employing sophisticated search and categorization methodologies and review work flows. Used in conjunction with TAR or appropriately on their own, these practices and tools may provide vast improvements over the traditional document-by-document or linear reviews, or even more modern simple "keyword" searches.

What are the Non-TAR Alternatives to Traditional Review or "Keyword" Searching?

Greer's publication starts from the premise that while advances in litigation technologies really do represent an improvement in results, their use (including TAR) in investigations has "not drawn much attention." These improvements can eliminate wide swaths of information through simple steps, and/or improve the odds of proper review and production for the subset of information actually reviewed.

Improvements that reduce the universe of reviewable information include the concepts of deduplication and near-deduplication, where identical (or near-identical) copies of documents are either eliminated within a data set, or grouped for more streamlined review. These practices can be used even across custodians, so that only one instance of a document is reviewed and subsequently produced.

Practitioners may also use objective and subjective criteria such as file extensions, search terms,

date or domain name restrictions to eliminate or cull files from the review population. Finally, sophisticated Boolean searches (combinations of keywords and operators such as "AND," "OR," and "[BUT] NOT") may be used across data sets to remove significant amounts of information.

Improvements that result in more efficient and targeted review and production incorporate mechanisms that cluster documents that seem to be similar (based on the frequency of certain words or custodians), which groups them together to allow reviewers to make consistent decisions about those groups of documents, rather than individual instances.

This type of grouping can also include email threading, which groups email messages together based on subject and participant. Additional grouping technologies may include Latent Semantic Indexing (LSI), a non-TAR algorithm that enables searches by concept rather than simple keywords. And as above, the same sophisticated Boolean searches that eliminate information can often find the information of highest value, shortcutting alternative, drawn-out investigations.

One further issue hinted at — but not discussed at length in Greer's paper — is the benefit of old-fashioned detective work, the correct starting point in any document collection and review process. A responding party should ask who is most likely to have relevant and important information, and whether there is a smaller group that the responding party should focus on first (after discussions with the division) rather than casting a wide electronic net over sources of information that will either be duplicative or less meaningful than richer sources of relevant information.

Each source of information added to the mix adds to the cost, burden and time associated with review. A responding party should determine whether there are reasonable bases to focus on some sources of data first, and add others only as necessary. But this approach requires transparency and meaningful discussions with regulators.

What About TAR?

TAR, a relatively new entrant into the world of litigation and investigations, is an iterative process through which human subject-matter experts (SMEs) interact with software and code small sets of documents. The computer takes into account the decisions of the subject-matter experts and generates new sets of documents from which it thinks it will learn from the human decision makers.

This process typically ends after a few thousand documents have been reviewed and the predictive coding tool concludes it can learn nothing more from the human reviewers. The predictive coding tool then extrapolates those judgments to the entire set of collected documents, and codes the documents as likely relevant or likely irrelevant.

This is not a "black box" or "set-it-and-forget-it" solution. Instead, the responding and requesting parties may need to (but not always) first agree on protocols covering how the system will be trained, when training will end and how the results will be audited.

The parties will likely also discuss how transparent the training process will be to the requesting party, and address questions such as, "Will the responding party share its relevance decisions during the training process?" and "How will the responding party handle privileged documents?" This may sound a bit more complicated than the traditional linear review, but TAR can provide efficiencies and consistency in return for that complication, particularly if it is handled in a manner that avoids disputes.

Further, TAR does not operate in a vacuum. For best results, organizations — and again, their attorneys — should be considering TAR as among their tools, rather than an "either/or" solution. Second requests and similar investigations can be complex, expensive and risky to the responding organization, making the effort practitioners expend developing a review road map time well spent.

These roadmaps start with the universe, and then utilize the proper tool or technology at the proper stage, often beginning with deduplication and other culling mechanisms to create a richer data set before employing additional means of search and review.

What are the Benefits of Improved Practices for Merging Parties Facing the Prospect of an Investigation?

Greer noted that the use of improved technologies — and, frankly, logic — offers "the promise of reducing the costs incurred by merging responding to second requests and the size the document productions received by the division, without undermining the ability of the division to conduct an appropriately thorough investigation."

Greer focused here on instances where the division "encouraged the merging parties to identify categories of information and documents responsive to the second request, but that are not particularly relevant to the dispositive issues in the investigation" where the "[p]roduction of these categories of information could be deferred or foregone entirely." (emphasis added)

With this in mind, and with apparent division blessing, organizations may employ clustering technologies that focus on information from the appropriate subset of individuals, rather than blindly following every lead or keyword "hit" within the population.

Greer offered several additional observations based on the division's negotiations of "TAR protocols in approximately a dozen instances." Based on that experience, Greer found that "TAR produced smaller, more responsive document productions," which "contained much more relevant information and less that obviously is not responsive." Greer also felt that the division staff benefited substantially and, based on reports from the responding parties, that the parties experienced "substantial time and cost savings" as well.

Greer went on to state that TAR provided additional opportunities to narrow party productions, including instances where the division "encouraged parties using a TAR protocol to identify categories of documents that, while technically responsive to the second request, [were] not essential to resolving the competitive concerns at issue in the investigation."

This focus on a protocol mirrors concerns raised by other regulatory or quasi-governmental bodies, such as the FDIC-R, which successfully complained of the use of TAR without a court-approved protocol modification.[1] But overall, rather than raising the protocol as a point, which would create contention between the parties, Greer instead presented the use of TAR as "an opportunity to reduce further the size of the production," which, in turn, saves the responding party money, and the responding party and the division time.

But Greer also included an important caveat when it came to the validation of a TAR process. That is, the division also consistently asked responding parties to "provide a statistically significant sample of nonresponsive documents to ensure that facially responsive documents were not excluded from the collection." Why? To support the use of TAR, the division was checking both the produced documents as well as samples of the data left behind, but the division did except "documents coded as privileged" from that nonresponsive review.

So What's the Takeaway?

Improvements in technology, appropriately applied, benefit the division by decreasing production volume and increasing production relevancy. These same steps benefit organizations by defensibly reducing the universe of reviewable and producible information, while responding to the regulators.

And while perhaps not precisely a win-win, this "benefit-benefit" should be preferable to the converse no one wants: a long, expensive, document-by-document, process in which electronically stored information is pared down using blunt instrument tactics, with only the certainties of bloated expense and delays in closing waiting at the end.

—By Jonathan L. Lewis, Gilbert S. Keteltas, Karin S. Jenson and James A. Sherer, BakerHostetler

Jonathan Lewis and Gil Keteltas are partners in BakerHostetler's Washington, D.C., office. Karin Jenson is a partner and James Sherer is counsel in the firm's New York office.

[1] Progressive Cas. Ins. Co. v. Delaney, No. 2:11-cv-00678-LRH-PAL, 2014 BL 140634 (D. Nev. May 20, 2014).