Published by Pike & Fischer, a BNA Company

Digital Discovery & e-Evidence

BEST PRACTICES & EVOLVING LAW

http://ddee.pf.com

Reprinted from Vol. 6, No. 7 | July 2006

Of Litigators and Butterflies: The Quest For A Quantum Leap in Large-Scale Document Review

By Nicolas Economou

You are the lead litigator in a high profile, high-stakes, document-intensive case. Imagine, if you will, an ideal document review process. Your review team consists of dozens of junior attorneys and paralegals, an unerring extension of your eyes and mind, able to assess at great speed and with perfect accuracy each of several millions of pages to determine their relevance to complex, evolving case issues.

Imagine if you could say: "It makes no difference if I review these documents for relevancy to the issues in my case, or if my review team does. We always agree on which documents are relevant, and which ones are not."

If you were fortunate enough to have this circumstance, you could then confidently peruse the subset of documents identified by your review team, and apply your legal expertise and case knowledge to pinpoint the few key documents that will support your client's legal claims.

But the reality is that an ideal large-scale document review process such as this does not exist — nor could it, no matter how capable the review team. Two questions naturally follow:

1. What prevents a perfect document review process from existing; and
2. What approach might get us (nearly) there.

This article endeavors to answer both of these questions.

Three Challenges to Document Review

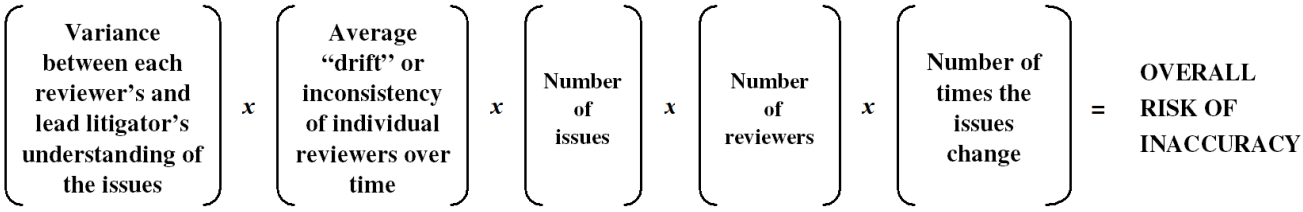
With respect to the first question, the three main obstacles to an ideal document review process are:

- Deficient Knowledge Transfer (a teaching challenge): It is impossible to ensure that individuals on any document review team consisting of more than a handful of reviewers will understand without any variance what the lead litigator is looking for. This deficiency becomes an even greater challenge when dealing with complex and changing case issues.
- Lack of Consistency (a human nature challenge): Even with perfect knowledge transfer, a consistently correct application of knowledge is simply impossible. Humans are prone to fatigue, loss of focus, distraction, snap judgments, and errors of reasoning or understanding. This challenge is magnified when combined with other variables such as complexity of subject matter, training, skills, dedication, workforce turnover, and time and cost constraints.
- Lack of Quality Assurance (a measurement challenge): In order to be effective, any document review process must be both self-evaluating and externally verifiable using widely accepted and rigorous quality assurance protocols. Absent the implementation of such protocols with which to evaluate the accuracy of review, the first two problems cannot be appraised, much less addressed.

Unless each of these three challenges is addressed, even the most well-designed and well-staffed large-scale document review process remains rife with risk.

The Risk Equation and the "Butterfly Effect"

Just how much risk is involved? Given the critical and pervasive nature of the three challenges discussed above, we can expect the amount of risk to be significant. The risk exposure can be estimated via a mathematical approximation. Factoring in the challenges discussed above, along with independent variables that come into play in any document review process, the overall risk of inaccurately assessing documents for relevance could be calculated as follows:



One could think of the risk so calculated as a manifestation of Chaos Theory's "Butterfly Effect" in document review: *Minor differences in initial conditions create disturbances that, over time, become immensely amplified and impossible to predict or control.*

Furthermore, even when the disturbances and inaccuracies become apparent, their size and scope are never precisely quantified in any way. This undermines not only the interests of the litigants, but also the integrity of the legal system itself. Clearly, an "ideal" review process cannot allow for this kind of risk.

Search Technologies in Context

Faced with these immense challenges and the potential risk they present, law firms and corporate legal departments often turn to keyword and "concept" search tools in hopes of containing risk and improving the efficacy of their review teams. However, the comfort that search tool technologies offer on their own is largely illusory for two reasons:

First, though search technology is clearly necessary and valuable in large-scale discovery, even expert users achieve only moderate levels of accuracy with current search tools. This is because keywords and "concepts" are generally poor indicators of relevance. Available academic data shows that even highly experienced reviewers using the most advanced search tools on a relatively small collection are likely to miss 50 percent or more of the documents that a human, upon review, would have assessed to be relevant[1].

Second, search software does not address the three major challenges identified above (knowledge transfer, consistency, and quality assurance). Search software cannot tell reviewers whether or not they clearly understand what the lead litigator is looking for, or notify them when they are inconsistent in their assessment of documents, or apprise attorneys of the actual performance of the overall review. In fact, as often as not, reliance on search software may cause reviewers to go faster in the wrong direction, and further amplify the Butterfly Effect throughout the document review system.

In Search of the Ideal

Turning now to the second question posed at the outset of this article, what sort of approach might get us closer to the "ideal" document review system? An alternative approach would need to do the following:

- Recognize that the three challenges discussed above — knowledge transfer, consistency and quality assurance — are largely non-technological process design challenges, and must be addressed as such.
- Recognize the need for experts: both process design experts to solve non-legal complex problems of the review system, as well as subject matter experts for each case.
- Recognize the proper functions and uses of search technologies in the overall document review process.
- Recognize the need to avoid disrupting business process — either the work of the attorneys and legal teams, or the firm's existing IT infrastructure.

Hence, the right approach to designing a markedly improved document review process would be conceptually "systemic" with respect to acknowledging and addressing the obstacles identified above, and it would be practicable. Because it would address the key underlying challenges of document review, we can expect that this type of approach will achieve higher performance.

Test Piece

Current evidence strongly supports the effectiveness of a systems-based approach as described above. In a controlled study, litigation management consultant Anne Kershaw evaluated H5's automated document review system. Described as a combination of professional services, process design expertise, and advanced information retrieval technology, H5's approach appeared to meet the criteria of being both conceptually systemic and practicable.

The question set forth in the study was: document for document, issue for issue, would the document review accurately replicate the litigators' complex relevancy assessments? Would its performance come anywhere close to that of the "ideal" document review process? How would it compare to the control process, a well-designed tradi-

tional manual review, conducted by a client's experienced reviewers with extensive case knowledge?

The salient finding of the evaluation was that H5's assessments of relevancy were nearly indistinguishable from the litigators'. Document for document, issue for issue, H5's review system correctly identified relevant documents as such 98 percent of the time — only two percent short of the elusive "ideal" review. The potential "disturbances" throughout the system were kept in check and, as a result, the Butterfly Effect had nearly disappeared.

By comparison, for each relevant document missed by the H5 system, the control review process missed 32 documents. That is, the risk of failing to flag relevant documents for litigator review was 32 times greater under the traditional review process.

The poor performance of the traditional review process was a clear manifestation of the Butterfly Effect: "noise" introduced by accumulated discrepancies in the understanding of the issues between the senior litigator and reviewers and by each reviewer's inconsistencies over time, became amplified through the system into a major disturbance.

Three Essential Components

To counter the three document review challenges outlined at the start of the article, it seems appropriate to conclude by proposing a three-part solution.

While there is more than one way to achieve an optimal document review, broadly speaking, an effective document review system must incorporate each of the following three components:

- A detailed definition of the target with which to make relevancy determinations. (Lists of keywords and concepts or a few exemplar documents do not constitute a reliable definition of what is being sought, and introduce considerable ambiguity into the process.)
- A review process in place (fully manual, fully automated, computer/search-assisted or some combination of these); and
- An evaluation protocol with which to measure performance.

To summarize: Based on a proper definition of the target, the results of the review — irrespective of whether it is conducted manually or with the support of a variety of technologies — can be assessed using reliable academic protocols.

As discussed earlier, legal practitioners have placed undue emphasis on the review process (second item, above), and in particular, on search tools. In doing so, they often overlook the other two components, which hold equal, if not greater, importance.

The first item (detailed definition of the target) is essential because if you don't determine what you're looking for, you can't measure what you've found. And the third item (measurement of performance) is crucial because if you don't measure your results, you can't improve them.

Arguably, this third item — utilization of performance evaluation protocols — holds the most importance in large-scale, complex document review. Yet surprisingly, this critical component is lacking in the overwhelming majority of document review processes today.

Reliable protocols, for example such as those utilized by TREC (Text Retrieval Conference) are widely available, relatively simple to implement, and can be applied at reasonable cost to measure accuracy. Cost and time-to-completion can also be measured reliably at the outset of a review, and hence help reduce overall risk.

If it is your prerogative to mitigate the risk of e-discovery mistakes, you first need to know your current risk. And in order to know your risk, you must measure the accuracy of your review system. By implementing performance evaluation protocols and processes by which you can measure — and improve — your performance, you will be able to manage your risk more effectively and keep the Butterfly Effect in check.

Endnotes

1. *See, e.g.*, the Text Retrieval Conference (TREC) studies, TREC-6 and TREC-7, http://trec.nist.gov/ (at acceptable levels of precision where at least 7 out of 10 documents are relevant, typically less than 50 percent of all relevant documents are retrieved). *See also* Turtle, 1994 (Westlaw study comparing manually constructed Boolean searches to automated system using natural language found that the Boolean approach returned only 24.4 percent of relevant documents, and the natural language system, though achieving superior results, retrieved just 32.9 of the relevant set).

Nicolas Economou *is chairman and CEO of* ***H5****, an e-discovery services and risk management firm. H5 advises Fortune 500 companies and AmLaw 200 law firms in the design and implementation of document review and analytic services for document-intensive litigation, records retention, and regulatory compliance.*

