April, 2014

Solving the High Cost of the "Review" Stage of Electronic Discovery

Ken Chasse, Mr.

Available at: https://works.bepress.com/ken_chasse/2/
Solving the High Cost of the “Review” Stage of Electronic Discovery

Ken Chasse

This article provides more details on the following Comment that I posted (April 10th) to Dan Pinnington’s article of April 8th, “Ontario Judge Strongly Pushes for Greater Use of Technology in Courts and Orders E-Trial”:

My Comment: “Make the preparation work of a lawyer making production comparable to that of an accountant. The client doesn’t give the accountant 100,000+ records and say, ‘here, you make up our financial records and then do the audit.’ The litigation lawyer should be able to work the same way, by combining the searching and reviewing into one act. That is how legal research is done, with the aid of headnotes, particularly their indexing. So, show the client how to make a comparable index of the client’s own records, which will give the client as much useful information for doing business daily, as do its financial records, as well as provide continuous preparation for e-discovery and trial. Then the high cost of the “review” stage of e-discovery will disappear. The lawyer would use the client’s index of records to prepare to make production—searching with the speed of electrons, instead of reading each record for relevance and privilege, or using a TAR device (technology assisted review device, such as uses “predictive coding”) which devices are based upon a faulty concept of e-review, and are still without a history of proven reliability. Again, the successful solution requires an adequate understanding of available technology. Learn it before you love it, whether it’s an electronic trial, an atomic bomb, or the consequences of striking a match or burning your toast.”

Searching an index is far faster, more precise, and therefore safer, than searching a whole database of all the texts themselves. Only a very small percentage of the words in texts are indexing words and phrases. And a lawyer’s searching an index for relevance and privilege is far safer than leaving the searching to the client. The lawyer should be enabled to do the searching by means of a good index of the client’s database of records. The client isn’t the legal expert in regard to the issues, which is the necessary expertise that determines the adequacy of the searching. But the client is an expert as to the nature and purpose of its business, which determines the adequacy of the indexing. That which is best done, is that which is expertly done.1

The concept upon which TAR devices are based is faulty, therefore using TAR is risky, because:

(1) it requires “reading” the whole of each record to determine its relevance, and privileged status;

(2) it doesn’t use the faster and more precise method of searching only indexing words and phrases, which is all that an index is made up of—only very small percentages of the words in texts, made up of sentences and paragraphs, are indexing terms;
(3) it is based on the faulty assumption that, because of the speed of electronic searching, indexing is no longer necessary; in fact, electronic searching of a good index is the fastest, most precise, and therefore the safest type of searching; that is what continues to justify the adding of headnotes and the indexing they start with, to judgements for publishing them;

(4) TAR devices are expensive, meaning big and high reward cases are needed to justify their cost;\(^2\)

(5) counsel needs a good “clawback agreement,” covering \textit{inter alia,} mistakes by TAR leading to the inadvertent production of privileged records. The “clawback agreement” enables arguing retrieval without adverse consequences, except for what’s left in the mind of one’s opponent who has read the records. But poor use of TAR might be equated to a waiver of privilege, as could inadvertent disclosure caused by poor records management. There should be a Canadian equivalent of Rule 502(b) of the U.S. \textit{Restyled Federal Rules of Evidence}, (effective from Dec. 1, 2011), which provides for such agreements. It states:

\begin{quote}
Rule 502. Attorney-Client Privilege and Work Product; Limitations on Waiver

The following provisions apply, in the circumstances set out, to disclosure of a communication or information covered by the attorney-client privilege or work-product protection.

\ldots

(b) Inadvertent Disclosure. When made in a federal proceeding or to a federal office or agency, the disclosure does not operate as a waiver in a federal or state proceeding if:

(1) the disclosure is inadvertent;
(2) the holder of the privilege or protection took reasonable steps to prevent disclosure; and
(3) the holder promptly took reasonable steps to rectify the error, including (if applicable) following Federal Rule of Civil Procedure 26(b)(5)(B).
\end{quote}

(6) TAR is sold to facilitate “back-end” searching by indexing; but in fact “front-end” indexing is far more accurate and economical. Indexing records when they are created or received is “front-end” indexing, which makes searching them for relevance and privilege later, much more efficient than indexing only when the searching is done at the “back-end” by way of the searching done for discovery purposes. Indexing is not just a mechanical process, but rather, requires thinking about the meaning and impact of what is being indexed, and therefore is best done when texts are made and received.

(7) records are written to convey information and not to facilitate indexing, sorting, or classification; therefore they may not contain any of the necessary terms for finding them by way of searching for standard indexing terms for their subject matter.
And good indexes contain those very helpful “see” and “see also” references that greatly increase the probability of finding all that is relevant. Texts don’t serve that purpose.

These last two reasons are especially true of older texts, written with the verbiage and assigned the indexing of an earlier time; e.g., things that are now, “cool” and “sexy,” were not so long ago, “neat” and “glamorous,” and if judicially quoted, should be thus represented in well indexed headnotes. So a searcher or indexer may have to translate his/her currently common language in order to adequately search for or index older texts and improperly motivated and misused language. 

But in spite of TAR’s serious shortcomings, compelled by the high cost of “review,” U.S. judges are now approving the use of TAR devices with which to allege adequate searching and production, even though such devices do not yet have a reputation for, let alone a history of, reliability. Such “rush to judgment” bespeaks a greater concern for limiting the cost of legal proceedings than for the consequential limiting of justice.

Searching the whole of a text requires TAR devices or human readers, to read sentences, paragraphs, and punctuation, among which, indexing terms are but a very small minority of the words and phrases. And all of that textual verbiage can be misleading as to what the subject matter is, especially so if the writing quality is poor. Indexing terms are written for purposes of searching and finding, therefore they convey only their concepts, issues, and related facts, and not literary devices such as, grammatical construction, emphasis, contrast, examples, analogies, drama, and the competence, experience, and integrity of the writer. Therefore, almost all of the words in records written in sentences and paragraphs, get in the way of searching for relevance and privilege, and anything else.

The Sedona text, The Sedona Canada Principles—Addressing Electronic Discovery (January 2008), approves of searching by “automated methods,” stating (p. 4):

Counterbalancing the dispersed nature of electronically stored information is the fact that some forms and media can be searched quickly and fairly accurately by automated methods. For these types of electronically stored information, lawyers may be able to search through far more documents than they could hope to review manually.

But it is very doubtful that TAR was known of by anybody as early as 2008.

TAR and the proportionality doctrine can be mutually interdependent for reducing the time and cost of e-discovery proceedings. But indexed databases are a better strategy, as is shown by all of the online services for legal research.

However, for sufficiently large databases, TAR devices could be used more safely for indexing at the front-end, when texts are created and received, or being prepared for publication, as are the indexing strings on the top of headnotes written for decisions from the courts. When clients start to build an index of previously created or received documents, TAR might be used, with frequent checking for accuracy.
When the indexing is caught up, thereafter the documents should be indexed as they are created or received. Then TAR shouldn’t be needed ever again.

Instructing clients on the use of indexing for database creation, use, and control, is part of the work of the “records management lawyer,” as is the work of electronic discovery specialists, which the bigger law firms are now appointing. As to other aspects of the work of such specialized lawyers I’ll save that for another article, because “records management law” will have to be a major field of the practice of law.

1 Therefore for example, never let the publisher write the index for the books that you write.

2 See: Christine Taylor, (technology industry consultant) “A call to vendors to make e-discovery affordable for ‘The Other 85 Percent’”; online: http://www.aceds.org/a-call-to-vendors-to-make-ediscovery-affordable/. It states in part: “The majority of e-discovery software platforms are made for big-budget cases. They serve millions of files, multiple review teams, hundreds of custodians and ESI collection across acres of disk and tape storage – and they get the lion’s share of media attention. A few years ago, these small firms performed e-discovery manually. But the e-discovery landscape is changing fast: rules are tightening and searchable data is growing. Modern e-discovery is fraught with the risks of irrelevant data, poor custodial practices, evidence spoliation, bloated processing costs, and expensive review. The 85% of mainstream practitioners have as much responsibility to avoid these mistakes as the 15%. But the lack of affordable ediscovery tools leaves small law firms scrambling for less-than-attractive options…”

3 For example, headnotes and indexing may require the untruthful use of the words such as “truth” and “truthful.” The Truth in Sentencing Act, S.C. 2009, c. 29 (operative from Feb. 22, 2010), purported to correct the “untruthful” sentencing practice of awarding 2 days reduction in the sentence imposed for each day spent in pre-sentencing custody (PSC), which was awarded in compensation for terrible jail conditions suffered during PSC, and lost eligibility for release on parole. But said the then Attorney General of Canada, Rob Nicholson, on Second Reading of the Act’s Bill C-25, on April 20, 2009, the real cause of jail overcrowding and related conditions, was the refusal of accused persons to apply for bail so as to obtain such “excessive” “2 for 1 credit.” The Act restricted such credit to 1:1, and to a maximum of 1.5 days for each day in PSC, “if the circumstances justify it” (Criminal Code s. 719(3),(3.1)). But in fact, no defence counsel ever had a client who said, “I didn’t do it, but, I don’t want you to apply for bail for me, because I want to get that “2 for 1 credit” if sentenced for what I didn’t do, regardless jail conditions, and the fact that I didn’t do what the police, the Crown, and the charges laid say I did, and even though you have told me that it is much more difficult to prepare a defence if I’m in jail.” See: Ken Chasse, “Untruth In Sentencing Credit for Pre-Sentence Custody” (2010), 15 Canadian Criminal Law Review 75. However, in R. v. Summers 2014 SCC 26 (April 11, 2014), the Court corrected such untruthful use of “untruth” in legislation by approving the use of such enhanced credit because of lost opportunity for early release and parole due to time spent in PSC not being taken into account. And the language of a “just sentence” used to include whipping with the “cat-o’-nine tails,” administered not less than ten days before the expiration of the related sentence of imprisonment; but females were not to be whipped (s. 668 of the 1970 Criminal Code; but repealed, S.C. 1972, c. 13, s. 59, thus surely saved from a successful Canadian Charter of Rights and Freedoms s. 15 “sexual equality” attack).


5 Ralph Losey, ibid., at 25, states that predictive coding (a variety of TAR) was slow to be used by U.S. lawyers until the decision of Magistrate Judge Andrew J. Peck on Feb. 24, 2012, in, Da Silva Moore v. Publicis Groupe 287 F.R.D. 182 (S.D.N.Y. 2012), approving the use of predictive coding, listing justifications. That decision was affirmed by a district court judge in, DaSilva Moore, 2012 WL 1446534 (S.D.N.Y.). Of similar effect was the decision of Judge Shira A. Scheindlin, whom Losey describes (ibid. p. 28) as, “the most influential judge in the e-

6 This is a main theme of the Losey article, supra note 4. It provides a detailed explanation as to how predictive coding works, and a detailed discussion as to its relationship to proportionality. However, this statement I think is premature, certainly for legal proceedings in Canada (ibid. at 8): “When used together, proportionality and predictive coding provide a viable, long-term solution to the problems and opportunities of the legal search of Big Data.”

7 See for example this announcement: “BLG hires Canada’s leading electronic discovery lawyer to lead e-discovery for the Firm. Toronto (August 6, 2013) — Borden Ladner Gervais LLP (BLG) is pleased to announce that renowned electronic discovery lawyer, Martin Felsky, will be joining BLG as National E-Discovery Counsel, effective August 12, 2013, to lead the Firm’s e-discovery and litigation support services. ….”

8 See for example this article: Ken Chasse, “Why a Legal Opinion is Necessary for Electronic Records Management Systems” (2012), 9 Digital Evidence and Electronic Signature Law Review 17, a U.K. “open source” journal, i.e., providing free downloading of articles (click “Archives” to access the contents of volume 9).