Manual Review and the Air Canada vs. West Jet Airlines Case

Marie-Andrée Vermette
WeirFoulds LLP

Presented at the E-Discovery Course
Toronto, February 20 & 21, 2008

The Manual Review Requirement

Marie-Andrée Vermette
WeirFoulds LLP

Presented at the E-Discovery Course
Toronto, February 20 & 21, 2008

The Manual Review Requirement

Contents

In the discovery process, the main purpose of document review is to identify the relevant documents to be produced and the privileged documents to be withheld. However, document review also plays a critical role in the broader context of the litigation: it is the stage where lawyers gain a greater understanding of the issues in the case and where legal strategies begin to develop. [1]

Document review has traditionally meant paper review. Such a review has now become impractical in many cases due to the dramatic increase in the volume of potentially relevant documents in the digital world compared to the paper world. In addition to eliminating the burden associated with printing all electronic documents for a paper review, an electronic review (i.e. the review of documents in electronic format) offers reviewers a greater ability to deal with significant volumes of electronic documents. [2] The use of electronic tools at the review stage can increase the efficiency of the review and assist in reducing the costs associated with document review. This is an important consideration given that the real cost of electronic discovery is generally not in the collection of the data, but in the human cost of reviewing the information. [3]

Given the volume of electronic data involved in certain cases, some have argued in favour of eliminating the step of document review in the e-discovery process. According to this view, after a party has performed data culling, such as the application of search terms, and done some quality control checks with respect to the culling techniques applied, that party's discovery obligations are fulfilled by

simply delivering the results of the culling and searches to the other parties. In other words, no "manual review" of the results, i.e. no review of the individual documents found to contain a "hit" or to match the culling criteria, would be required. This approach was recently rejected in *Air Canada v. WestJet Airlines Limited.* [4] In this case, which involved more than 75,000 potentially relevant documents, Nordheimer J. confirmed the requirement that a manual review be performed after the application of search terms.

Despite its apparent burdensome character, conducting a manual review can be very beneficial for the case and the lawyers involved, and can be made easier by the use of electronic tools and techniques. In this paper, I will first review briefly the step preceding the review process, i.e. data culling. I will then examine the requirement to conduct a manual review, and discuss how to make this task easier.

I. Preliminary Step: Data Culling

The manual review requirement does not apply to all the electronic data of a party, or even to all the documents contained in a data set identified as potentially including relevant information (e.g. all e-mails sent and received by the employees working in a particular department). This is because large collections of data typically contain vast amounts of duplicate and irrelevant information. The objective of data culling (also known as data screening or filtering) is to reduce the data set to a more manageable set of potentially relevant documents for

review. The effectiveness of the culling strategies employed can have a significant effect on the costs related to processing and review: in large data collections, culling rates often exceed 85%. [5] However, the strategies adopted must be reasonable and defensible in order to achieve the desired results and to withstand the scrutiny of the Court (and the opposing party). Some culling techniques – such as de-duplication - are widely used and accepted, but attention must still be paid to the manner in which they are applied and to their ramifications. The reasonableness of other culling techniques – such as the application of search terms – will depend on the specific culling criteria applied in each case and whether they are adequate for the purpose of identifying the relevant documents within the overall collection of documents.

The issue of data culling is addressed in Principle 10 of the *Ontario Guidelines* which provides the following:

> A party may satisfy its obligation to produce relevant electronic documents in good faith by using electronic tools and processes, such as data sampling, searching, or the use of selection criteria, to identify the documents that are most likely to contain relevant data or information.

The Commentary that accompanies Principle 10 sheds some light on its meaning. It states:

> Particularly where searches for relevant electronic documents must be undertaken on large computer systems, containing vast amounts of information, including materials that are likely to be irrelevant, it may be impractical or prohibitively expensive to review all that information for relevance and privilege. In such circumstances, it is reasonable for parties to use electronic

techniques to search within electronic document sources, in collecting the materials that will be subject to a detailed review for relevance and privilege. **The objective should be to identify a subset or subsets of the available electronic documents for detailed review, that are most likely to be relevant.**

Where possible, parties and counsel should agree in advance on the search methods, and selection criteria or search terms, that will be used. Absent such agreement, however, parties should record and be prepared to disclose any limits on the searches they have undertaken, and to outline the scope of what they are producing, and what potential sources or documents have not been searched. [6] [Emphasis added.]

The approach adopted by the *Ontario Guidelines* regarding data culling is the same as the one reflected in the *Sedona Principles*. Principle 10 of the *Ontario Guidelines* is, for all intents and purposes, identical to Principle 11 of the *Sedona Principles*. [7] The *Sedona Canada Principles*, in particular Principle 7, also adopt the same approach. [8]

The application of search terms or keywords is one of the data culling techniques that is the most commonly used. When using search terms at the culling stage, the objective is to narrow the data set to include all relevant documents and segregate documents that are not responsive. However, it is impossible to realize this objective fully due to a number of limitations when conducting keyword searches or any other query-based searches (e.g. concept searches). It is important to review these limitations briefly because they explain in part why it is important to conduct a manual review.

The most important limitation is that, because of the lack of standardized terms used in conversations and documents, it is unlikely that keyword searches retrieve all relevant documents to the search terms employed. Electronic searching simply matches the word without any regard to the meaning of the keyword. As a result, the formulation of a query can be difficult if the keywords have numerous synonyms and can be described in numerous different ways. Even if the "right" keywords are used, queries often locate ambiguous uses of the keywords and retrieve "hits" of the words that are not relevant. In addition, a keyword search will not retrieve documents containing a keyword if the keyword is misspelled in the query or in the documents. [9]

Two of the concepts used to discuss the ability of keyword searching to retrieve the information that is relevant are recall and precision. The recall rate is the number of relevant documents retrieved compared to the total number of relevant documents in the data set. Thus, if an information retrieval technique achieves 70% recall, this means that 70% of all relevant documents were actually found and 30% of all relevant documents were not found. The precision rate, for its part, only relates to the retrieved documents, and represents the number of relevant documents retrieved compared to the total number of documents retrieved. Thus, if a system has a precision rate of 65%, 65% of all the documents retrieved were actually relevant, and 35% of the documents retrieved were not relevant. [10]

These two concepts can be further illustrated by an example. If, in a database of 1,000 documents, there are 100 relevant documents, and the search methodology used retrieves 200 documents, with only 50 of these 200 "hits" being relevant, the recall rate is 50% (50/100), and the precision rate is 25% (50/200). While precision is an easy percentage to measure, recall is usually not because it would require the person conducting the search to know the total number of relevant documents in the database and compare that number with the number of relevant documents found using a particular search query. [11]

Recall and precision are inversely related. Consequently, trying to improve the performance of one factor will cause the performance of the other factor to decrease. This is illustrated by the fact that if one broadens the scope of the search to increase recall (i.e. the number of relevant documents captured by the search), the recall will increase but the precision will decrease because such a search will retrieve more documents, many of them irrelevant. Conversely, if one wants to increase precision and adds keywords to the search query to limit the number of documents retrieved, precision will increase but recall will decrease, possibly omitting a number of highly relevant documents. [12]

Search features typically support either recall or precision. For example, the use of proximity connectors increases precision but decreases recall, and wildcard searches increase recall but decrease precision. [13]

Manually-created Boolean queries have generally been found to perform poorly with respect to both recall and precision in a number of studies. The two key limitations of this type of queries is that it requires negotiation of a very steep trade-off between recall and precision, and it leaves the user in ignorance as to the actual performance of the system. These limitations also apply to other types of query-based methods, such as natural language systems and the "concept searches" built on top of them, which have also been found to have poor results in terms of recall and precision. According to some studies, query-based methods can be expected to retrieve no more than 50%, and in all probability much less than 50%, of the documents they were employed to find, and will do so at the cost of including a very large amount of irrelevant documents in the "hits". [14]

Further, the fact that recall and precision are inversely related poses a dilemma to lawyers. On the one hand, recall is critical in litigation because parties have an obligation to produce all relevant documents, and a more complete set of relevant documents reduces the risk of overlooking information that could be critical to the case. On the other hand, precision is also very important because a low precision rate means that a large number of irrelevant documents will have to be reviewed, resulting in considerable costs and delay. [15] Nevertheless, it is essential to recognize this issue and not to succumb to the "false confidence syndrome" which has been documented by a number of studies and occurs when a relatively high precision leads the person conducting the search to overestimate recall and be falsely confident that the research results are satisfactory. [16] In one of these

studies, the Boolean queries that had been performed had a reasonably good precision (79%), but a low recall (20%). Despite the fact that they had succeeded in retrieving only one out of every five relevant documents, the lawyers involved believed from their evaluation of the retrieved documents that they had found three out of every four relevant documents. [17]

Should we conclude from the foregoing that the use of search terms is not an acceptable technique to exclude documents from the review set, and that the only acceptable practice to locate relevant documents is a full manual review of the entire data set? Such a position would be untenable and unrealistic in many cases given the volume of electronic data involved. Further, this position ignores the weaknesses and challenges associated with a review performed exclusively by human reviewers, and the fact that such a review will not locate all the relevant documents in the data set. Insisting on a full manual review, at any cost, for the purpose of obtaining all documents that are possibly relevant, no matter the degree of relevance, would also go against the guidelines and principles referred to above in relation to data culling, as well as two principles that inform all the e-discovery guidelines that have been developed so far, i.e. the principle of reasonableness and the principle of proportionality. [18]

The fact that the use of quey-based searches is now inescapable, however, does not mean that lawyers should ignore their inherent limitations. Without a proper understanding of the limitations of search and filtering technologies, lawyers will

be lulled into believing that they have completed a comprehensive search for relevant electronic documents and will fail to take appropriate steps that could locate additional relevant materials. This could result in an incomplete record that makes it difficult to prove certain aspects of the case, or could provide the basis for motions to compel production and allegations of spoliation. Lawyers who have a good knowledge and awareness of the limitations of query-based searches will be able to develop a better search methodology. Further, they will be in a better position to attack their opponents' methodology and to ensure that such methodology is adequate and reasonable in the circumstances. What the limitations associated with keyword searches emphasize, though, is the importance of carefully developing one's list of search terms, of adopting an iterative process (as discussed further below), and of implementing quality assurance mechanisms to test the searching strategies adopted and to measure their performance. All of these steps require, at a minimum, that a manual review be conducted of the documents identified through the application of search terms. Subjecting the entire mailboxes of key witnesses to a message-by-message manual review is another "mechanism" or process that is recommended in order to address the problem of the lack of standardization of language and the weaknesses associated with keyword searches. [19]

In light of the limitations inherent in query-based searches, conducting a manual review after the application of search terms is highly desirable. It is now also a legal requirement in almost all cases.

II. Necessary Step: Manual Review

The application of filtering criteria, search terms or advanced searching techniques such as concept searching does not eliminate the need for a manual review.

The need for a manual review after the use of searching tools and techniques is recognized by the *Ontario Guidelines* which provide in the Commentary that follows Principle 10 (reproduced above) that the objective of electronic searching techniques is to identify a subset or subsets of the available electronic documents that are most likely to be relevant for detailed review. Thus, although the use of searching techniques can significantly decrease the number of documents to review, it does not eliminate the need to review electronic documents altogether. [20] The *Ontario Guidelines* also note that over-production of irrelevant electronic documents – which is bound to happen in the absence of a manual review -- may be just as damaging to the client's interests and the litigation process as incomplete production. [21]

Principle 11 of the *Sedona Principles* is to the same effect as Principle 10 of the *Ontario Guidelines* in that Comment 11.a states that search methodologies allow the identification of potentially responsive information, which must then be reviewed to satisfy the party's search obligations. [22] In contrast to the *Ontario Guidelines* and the *Sedona Principles*, the *Sedona Canada Principles* contain extensive remarks that purport to deal specifically with the issue of review of

electronic documents in Comment 7.c, including the statement that "[r]esearch in the information science field has demonstrated that automated review is statistically more reliable than human review of large data collections for the purpose of identifying relevant electronically stored information". [23] However, the American authority referred to in support of this statement and the text of Comment 7.c in general relate to search and retrieval methods, i.e. data culling, not to the review of the documents identified as a result of the application of search terms and other culling techniques. In fact, the American article referred to in Comment 7.c expressly recognizes the importance of performing a manual review of the search results for relevance and privilege. [24] Another odd aspect of Comment 7.c is that it almost totally ignores the seminal Canadian case on the issue of review, *Air Canada v. WestJet Airlines Ltd.*, [25] referring to it only in passing in a footnote and without discussing its clear ruling. In light of the above, it is difficult to distil the meaning of Comment 7.c and one hopes this comment will be clarified in future versions of the *Sedona Canada Principles*.

As stated above, the need for a manual review to be performed by lawyers has also been recognized in the case law, notably in *Air Canada v. WestJet Airlines Ltd.* [26] In that case, Air Canada had used search terms to search its electronic database for relevant documents, and had then subjected the documents thus identified to an electronic filter intended to flag privileged documents. Air Canada was proposing to produce the resulting electronic documents without any further review for relevance, privilege or confidentiality, except for a sample review of

less than 5% of the relevant documents, the purpose of which was to ensure that the electronic search was effective. [27] WestJet Airlines Ltd. objected to Air Canada's proposed way to proceed. Nordheimer J. summarized WestJet's position as follows:

> The principle [sic] dispute between the parties is over the intention of Air Canada not to undertake a manual review of the documents to be produced. WestJet says that a manual review of all of the documents to be produced is necessary both because that is the obligation of a party under the *Rules of Civil Procedure* but also because it is the only effective way of determining whether the documents to be produced are relevant. Given the different meanings that words can have, electronic searches alone cannot be relied upon by themselves to distinguish between documents that use a word in a relevant context over documents that use the same word in an irrelevant context. WestJet also says that, given the importance of solicitor client privilege in our justice system, it is inappropriate for a party not to take all available steps to ensure that privileged documents are not produced nor is it appropriate for a court to countenance that failure by giving a blanket order of the type sought by Air Canada. Further, WestJet says that no electronic search can determine whether a document is properly labelled confidential. A manual review of each of the documents must be undertaken before Air Canada can fairly label a document as Level A or Level B. Simply put, WestJet says that what Air Canada is actually trying to do through this motion is to foist onto the defendants, Air Canada's obligations to identify and produce only relevant and non-privileged documents. [28]

Nordheimer J. agreed with WestJet's position. After referring to Principle 10 of the *Ontario Guidelines* and the Commentary following that principle, Nordheimer J. stated as follows:

Air Canada says that its proposed manner of proceeding is consistent with Principle #10 and should be endorsed by this court. I do not agree. I accept that the first stage of Air Canada's approach was appropriate, that is, the use of electronic search terms to identify the apparently relevant documents. WestJet does not dispute this. I do not accept, however, that Air Canada's intention not to conduct a manual review of the resulting documents is validated by Principle #10 nor is it consistent with the requirements of the *Rules of Civil Procedure*.

In my view, it is clear from the Commentary to Principle #10 that some form of further review is contemplated after the electronic search has been completed. The Commentary expressly refers to a "detailed review for relevance and privilege". While the Commentary does not expressly say that such a detailed review must be a manual review, and while I am prepared to accept that in some cases such a detailed review might possibly be conducted electronically, in the circumstances of this case, I do not see how that detailed review could properly be accomplished other than manually. I agree with counsel for WestJet that electronic searches alone cannot distinguish between documents that use a word in a relevant context over documents that use the same word in an irrelevant context since words have different meanings in different circumstances. For example, one only has to look at a search term such as "capacity" to realize that it could be found in an operational document dealing with passenger capacity and also be found in a human resource document dealing with the capacity of an individual to perform his or her duties.

I also agree with counsel for WestJet that solicitor and client privilege is too important a principle for the court to approve of a process that, on its face, has a very large potential for the disclosure of privileged material. Solicitor and client privilege is fundamental to the justice system in Canada [...]. It is not a principle that should readily be sacrificed to the interests of expediency or economics.

[...]

Having said that, it does not follow from my conclusions that each and every page of each and every document was [sic] be manually reviewed. Presumably different categories of documents will require different levels of review. It is up to Air Canada and its counsel to determine to what extent a detailed review of the electronic documents must be conducted. They must do so, however, cognizant of the obligations under the *Rules of Civil Procedure* regarding the production of documents that include the requirement that an officer of the corporation swear an affidavit of documents that confirms that a "diligent search" has been conducted for all documents. [29]

In his reasons, Nordheimer J. stated that he was prepared to accept that in some cases the review of search results may not require a manual review and could possibly be conducted electronically. However, he did not give any indication as to what type of cases would not require a manual review. In light of Nordheimer J.'s comments with respect to privilege, the limitations inherent in the use of search terms and the requirements of the *Rules of Procedure*, and in light of the fact that Air Canada was ordered to conduct a manual review despite the very high number of documents involved, it is difficult to imagine a case where Nordheimer J. would not require that a manual review be performed. Thus, the exception to the manual review requirement appears to be very narrow, but it will probably be clarified as the case law develops.

As found by Nordheimer J., Canadian discovery rules also support the requirement that a manual review be performed. In Canada, contrary to the United States, there is a general duty to produce relevant documents imposed by the different provincial rules of civil procedure. [30] These rules typically require the

preparation and delivery of an affidavit or list of documents. [31] In fulfilling their discovery obligations under the rules of civil procedure, both the parties and their counsel have certain obligations. In Ontario, for example, parties must swear or solemnly affirm that: (1) they have conducted a diligent search of their records and have made appropriate enquiries of others in order to make the affidavit of documents; and (2) the affidavit discloses, to the full extent of the party's knowledge, information and belief, all documents relating to any matter in issue in the action that are or have been in the party's possession, control or power. [32] In order to be in a position to make these statements in their affidavits of documents, parties must conduct a manual review to ensure, among other things, that the search terms that were used were effective, and that certain areas of relevance or certain custodians were not overlooked.

The Ontario *Rules of Civil Procedure* also require a party's lawyer to certify on the affidavit of documents that he or she has explained to the deponent: (a) the necessity of making full disclosure of all documents relating to any matter in issue in the action; and (b) what kinds of documents are likely to be relevant to the allegations made in the pleadings. [33] Apart from this express legal requirement, it has been recognized for a long time that counsel have certain duties and responsibilities in connection with the discovery of documents. In fact, it has often been stated in the case law that nowhere in civil procedure is the responsibility of lawyers greater than in the area of discovery of documents. Among other things, counsel must advise their clients as to what is involved in the

discovery process and must ensure that their clients disclose all the documents that may have a bearing on the case. Further, counsel have the responsibility to the opposing party and to the court to see that the affidavit of documents is confined to documents that may have a bearing on the case, and that it adequately and accurately describes the documents. [34] Again, these duties cannot be fulfilled by the mere application of search terms, without a subsequent manual review of the results.

The general practice in the United States also includes a manual review for relevance and privilege after the application of search terms. [35] This is consistent with the obligations that have been imposed on counsel in the United States, which are similar to counsel's obligations in Canada. In *Zubulake v. UBS Warburg LLC*, Scheindlin J. stated that counsel have a duty to monitor their clients' efforts to retain and produce relevant documents, to make certain that all sources of potentially relevant documents are identified and preserved, and to produce information responsive to the opposing party's requests. [36] Similarly, in *Cardenas v. Dorel Juvenile Group, Inc.*, Waxse J. stated that counsel have an obligation to review the documents received from their clients to see whether they indicate the existence of other documents not previously retrieved or produced. This obligation flows from counsel's duty to ensure that their clients discharge in good faith their duties under the discovery provisions of the *Federal Rules of Civil Procedure*, and counsel's duty to exercise some degree of oversight over their clients to ensure that they are acting competently, diligently and ethically in

order to fulfill their responsibility to the court and opposing parties. Waxse J. held that counsel had not met these duties in that particular case, and found that the imposition of monetary sanctions was appropriate. [37]

In addition to the rules and principles referred to above, the limitations that are inherent in the use of search terms, discussed in the preceding section of this paper, and the fact that discovery is an iterative process support the requirement of manual review and show its importance.

Like paper discovery, electronic discovery requires an iterative process. The importance of adopting such a process has been described as follows:

> It is critical that electronic discovery be viewed as an iterative process. While the initial identification phase may have turned up five custodians of interest, a handful of keywords and concepts of interest, and a few months in which company events are of interest, analysis typically expands this set of criteria throughout the course of the investigation. Simply performing search and collection once or twice up front seldom proves sufficient, and further collections are generally warranted throughout the discovery effort. Assumptions concerning date ranges made early in the analysis phase can be wrong, in that certain key dates may have been unknown or incorrect. Assumptions made concerning keywords, phrases and concepts of interest may have been incomplete. As analysis yields results, new keywords, phrases and concepts, and in some cases new custodians of interest are added to the scope of the investigation, requiring re-searching of the original (or a subset thereof) data set.
>
> Failure to employ an iterative process can lead to an incomplete set of evidence which could, of course, jeopardize your case. [38]

In short, cases evolve. In many instances, the issues that were thought to be important at the beginning of the litigation are different from those deemed to be important later on. The identification of new issues through the review and analysis of the collected data set frequently reveals the need to collect additional documents or to modify one's list of search terms and filtering criteria because of newly discovered relevant custodians, topics, time frames, or keywords. [39] The *Sedona Principles* recognize the importance of following an iterative process, and state that agreements between parties on the search methods to be used, including search terms and concepts, "should take account of the iterative nature of the discovery process and allow for refinement as the parties' understanding of the relevant issues develop." [40]

A necessary ingredient of an iterative process is a manual review performed by human reviewers. The mere application of search terms or advanced searching techniques such as concept searching will not, without a manual review of the results, reveal the need to collect additional documents, or the need to apply different search terms because of the particular jargon used by the key players. [41] As stated in *The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery*:

> Using a search and retrieval technology in conjunction with an implementing process in the complex context of electronic discovery will involve multiple phases of activity, with iterative feedback opportunities at appropriate decision points to allow integration of what a case team learns after each exercise of the process in order to calibrate and maximize the technology's

capability to identify relevant information. It is through this feedback that case teams will acquire sound information to use in making both strategic and tactical decisions.

The initial search and retrieval process should be designed with the intent that it serve as a pilot process that can be evaluated and modified as the team learns more about the corpus of information to be reviewed. One useful approach is to initiate the process by focusing on the information collected from a few of the custodians who were at the center of the facts at issue in the litigation or investigation. Focusing on information collected from the core custodians, which has a higher likelihood of being relevant, will help the team efficiently develop its understanding of the issues and language used by the custodians, thus allowing them to more efficiently develop and implement an appropriate search and retrieval process. [42]

Further, not reviewing search results before production only postpones the time at which the documents will have to be reviewed. Chances are that counsel will want to review the documents that his or her client produced to the opposing party at some point. Reviewing these documents earlier in the process can assist the legal team in its assessment of the case, in preparing for the next steps in the litigation, and in developing and testing theories of the case.

Thus, conducting a manual review of search results is not only required by law in most cases, it also makes sense. Conducting such a review can be, in some cases, very time-consuming and burdensome, but there are ways to make this task easier.

III. Making the Manual Review Process Easier

The use of various searching and grouping tools can assist in conducting a manual review. Search techniques can be used to perform an initial assessment of your

case, and to prioritize and organize the review of electronic documents. These techniques include "traditional" techniques such as keyword searching, and emerging technologies such as concept and context searching, clustering, guided navigation and visualisation.

The use of searching and grouping tools can be highly beneficial to the review process. Reviewing similar documents together can significantly speed up the review, by as much as ten times according to some. This is because knowing that the documents in topic folders are similar allows the reviewer to categorize the documents faster. Further, grouping documents together can allow for the bulk identification of key documents, the isolation of non-responsive data, and the identification of documents that need to be reviewed separately from the general document collection. This can be of great assistance in determining review assignments and resource allocation. Thus, groups of documents that are potentially privileged can be assigned to more senior reviewers, while groups of documents that are likely to be irrelevant can be assigned to first-year associates. Similarly, categories of documents like spam mail may potentially be ignored altogether or assigned to junior reviewers, and foreign language documents can be separated and assigned to reviewers with the appropriate expertise. [43]

Another advantage of these tools is to give the reviewers a more global view of the data set:

Getting a "wide perspective" of the electronic document universe is another highly beneficial aspects of emerging technologies like linguistic patterns and conceptual searching. Lack of structure is one of the biggest challenges when dealing with electronic documents and electronic discovery. When contextual similar documents are grouped together, certain facts about the document universe may, [sic] be discerned. For instance, a grouping of like documents will often result in threads of e-mail messages sent, received, replied to and forwarded among a group of custodians. A reviewer can see the full context of the electronic exchange and make a more informed categorization decision regarding the group of e-mails as a whole, rather than categorizing each message as a stand-alone communication. [44]

As suggested in the excerpt above, searching and grouping technologies allow for categorization consistency. Grouping like documents together and having them reviewed by the same person can significantly assist in categorizing in a consistent fashion documents with similar issues of relevance, documents with similar issues of confidentiality, and privileged documents. [45] In addition, grouping tools can provide a solution to the problem of near duplicates. When near duplicates are grouped together at the outset of the review, batch analysis and coding of documents is greatly facilitated, thereby ensuring consistency and saving reviewer time and costs. [46]

Moreover, the use of searching and grouping tools can often lead to the identification of new issues or unknown items related to the case which require further investigation. It can also assist in identifying patterns, relationships or connections that reveal critical details of the case. [47]

Finally, given that many searching and grouping technologies also rank documents by relevance to the searched keyword or concept, reviewers can prioritize their review time and concentrate first on the documents that are most likely relevant. [48]

IV. CONCLUSION

Although the rules applicable to "paper" discovery often have to be adapted to the world of electronic discovery, such adaptation does not require the complete elimination of the manual review stage in the e-discovery process. This is because conducting a manual review after the application of culling criteria is necessary in order to fulfil the objectives of the discovery process and to meet the parties' obligations in that process. This is likely to be the case as long as the "goal of computational thinking to approximate the ability of human language behaviour" remains unfulfilled. [49] However, while we are not at a point where fully automated review can replace human review, advanced technologies can be used to focus review efforts on those documents which are most likely to contain relevant information, thereby increasing efficiency, reducing costs and providing superior results.

Endnotes

1. Electronic Discovery Reference Model, "Review Node", available at www.edrm.net/wiki/index.php/Review_Node.

2. Electronic Discovery Reference Model, "Review – Evolution of Document Review", available at www.edrm.net/wiki/index.php/Review_-_Evolution_of_Document_Review.

3. M.R. Arkfeld, *Electronic Discovery and Evidence*, 2006-2007 ed. (Phoenix: Law Partner Publishing, 2006) at p. 5-51.

4. (2006), 81 O.R. (3d) 48 (S.C.J.).

5. M. Mack and M. Deniston, *A Process of Illumination: The Practical Guide to Electronic Discovery* (Portland, Oregon: Discovery Center for Excellence, 2004) at p. 70; Electronic Discovery Reference Model, "Processing – Scoping Electronic Discovery Projects", available at www.edrm.net/wiki/index.php/Processing_-_Scoping_Electronic_Discovery_Projects.

6. *Guidelines for the Discovery of Electronic Documents in Ontario*, at pp. 14-15, available at http://www.oba.org/en/pdf_newsletter/E-DiscoveryGuidelines.pdf.

7. *The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production*, 2nd ed., June 2007 (Sedona, Arizona: The Sedona Conference, 2007) at p. 57, available at www.thesedonaconference.org.

8. *The Sedona Canada Principles: Addressing Electronic Discovery*, January 2008 (Sedona, Arizona: The Sedona Conference, 2008) at p. 27, available at www.thesedonaconference.org.

9. M.R. Arkfeld, *Electronic Discovery and Evidence*, 2006-2007 ed. (Phoenix: Law Partner Publishing, 2006) at pp. 5-19 – 5-20, 5-21.

10. *Ibid.* at pp. 5-20, 5-25; C.M. Bast & R.C. Pyle, *Legal Research in the Computer Age: A Paradigm Shift?* (2001) 93 Law Library Journal 285 at pp. 292-293; J.R. Baron, "Toward a Federal Benchmarking Standard for Evaluating Information Retrieval Products Used in E-Discovery" (2005) 6 Sedona Conf. J. 237; H5, "Information Retrieval 101: A Primer on Precision and Recall", 2006, available at http://www.h5technologies.com/pdf/IR101.pdf.

11. M.R. Arkfeld, *Electronic Discovery and Evidence*, 2006-2007 ed. (Phoenix: Law Partner Publishing, 2006) at pp. 5-20, 5-25; C.M. Bast & R.C. Pyle, *Legal Research in the Computer Age: A Paradigm Shift?* (2001) 93 Law Library Journal 285 at pp. 292-293.

12. M.R. Arkfeld, *Electronic Discovery and Evidence*, 2006-2007 ed. (Phoenix: Law Partner Publishing, 2006) at p. 5-20, 5-25; C.M. Bast & R.C. Pyle, *Legal Research in the Computer Age: A Paradigm Shift?* (2001) 93 Law Library Journal 285 at pp. 292-293; J.R. Baron, "Toward a Federal Benchmarking Standard for Evaluating Information Retrieval Products Used in E-Discovery" (2005) 6 Sedona Conf. J. 237.

13. M.R. Arkfeld, *Electronic Discovery and Evidence*, 2006-2007 ed. (Phoenix: Law Partner Publishing, 2006) at p. 5-25.

14. B. Hedin, "Searching in all the Wrong Places: The Effectiveness of Search Tools in E-Discovery", January 2007, available at http://www.h5technologies.com/pdf/searchtools.pdf.

15. H5, "Information Retrieval 101: A Primer on Precision and Recall", 2006, available at http://www.h5technologies.com/pdf/IR101.pdf.

16. M.R. Arkfeld, *Electronic Discovery and Evidence*, 2006-2007 ed. (Phoenix: Law Partner Publishing, 2006) at p. 5-20 – 5-21; C.M. Bast & R.C. Pyle, *Legal Research in the Computer Age: A Paradigm Shift?* (2001) 93 Law Library Journal 285 at pp. 292-293; J.R. Baron, "Toward a Federal Benchmarking Standard for Evaluating Information Retrieval Products Used in E-Discovery" (2005) 6 Sedona Conf. J. 237.

17. B. Hedin, "Searching in all the Wrong Places: The Effectiveness of Search Tools in E-Discovery", January 2007, available at http://www.h5technologies.com/pdf/searchtools.pdf.

18. See, e.g. Principles 2 and 5 of the *Sedona Canada Principles*; Principles 2, 5 and 10 of the *Ontario Guidelines*; and Principles 2, 5 and 11 of the *Sedona Principles*.

19. C. Ball, "Discovery of E-Mail: The Path to Production", 2006, at p. 6, available at http://www.craigball.com/The%20Path%20to%20Production%20of%20E-Mail.pdf.

20. *Guidelines for the Discovery of Electronic Documents in Ontario*, at p. 15, available at http://www.oba.org/en/pdf_newsletter/E-DiscoveryGuidelines.pdf.

21. *Ibid.* at p. 6.

22. *The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production*, 2nd ed., June 2007 (Sedona, Arizona: The Sedona Conference, 2007) at p. 57, available at www.thesedonaconference.org.

23. *The Sedona Canada Principles: Addressing Electronic Discovery*, January 2008 (Sedona, Arizona: The Sedona Conference, 2008) at p. 29, available at www.thesedonaconference.org.

24. J.R. Baron, D.D. Lewis & D.W. Oard, "TREC-2006 Legal Track Overview" at p. 2, available at http://trec.nist.gov/pubs/trec15/t15_proceedings.html. It is also noteworthy that a more recent article involving one of the authors, also referred to in the *Sedona Canada Principles*, states that "the comparative efficacy of the results of manual review versus the results of alternative forms of automated methods of review remains very much an open matter of debate": see J.R. Baron, ed., "The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery" (2007) 8 The Sedona Conference Journal 189 at pp. 194, 199, available at www.thesedonaconference.org under "Publications". This article also recommends that the legal community support collaborative research with the scientific and academic sectors aimed at establishing the efficacy of a range of automated search and information retrieval (at p. 195).

25. (2006), 81 O.R. (3d) 48 (S.C.J.).

26. *Ibid.* See also *Farris v. Staubach*, 2006 CanLII 19456 at para. 22 (Ont. S.C.J.), and *Shell Canada Ltd. v. Superior Plus Inc.*, [2007] A.J. No. 1359 at paras. 30-35 (Q.B.) (QL).

27. *Air Canada v. WestJet Airlines Ltd.* (2006), 81 O.R. (3d) 48 at para. 7 (S.C.J.).

28. *Ibid.* at para. 10.

29. *Ibid.* at paras. 13-15, 17.

30. Except in Quebec, which has a different legal regime.

31. *The Sedona Canada Principles: Addressing Electronic Discovery*, January 2008 (Sedona, Arizona: The Sedona Conference, 2008) at p. 8, available at www.thesedonaconference.org.

32. Rule 30.03 and Forms 30A and 30B of the Ontario *Rules of Civil Procedure*, R.R.O. 1990, Reg. 194.

33. Rule 30.03(4) of the Ontario *Rules of Civil Procedure*, R.R.O. 1990, Reg. 194.

34. *Delta Electric Co. Ltd. v. Aetna Casualty Co. of Canada*, [1984] N.B.J. No. 112 at para. 22 (Q.B.) (QL); F.D. Cass, P.Y. Atkinson & J.J. Longo, *Discovery: Law, Practice and Procedure in Ontario* (Toronto: Carswell, 1993) at pp. 3-9.

35. See, e.g., *Medtronic Sofamor Danek, Inc. v. Michelson*, 2003 U.S. Dist. LEXIS 8587 at pp. 5-6, 32-39 (W.D.Tenn. 2003); M. Mack and M. Deniston, *A Process of Illumination: The Practical Guide to Electronic Discovery* (Portland, Oregon: Discovery Center for Excellence, 2004) at p. 85-86; J.R. Baron, ed., "The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery" (2007) 8 The Sedona Conference

Journal 189 at p. 209, available at www.thesedonaconference.org under "Publications".

36. 229 F.R.D. 422 at pp. 432-433 (S.D.N.Y. 2004).

37. 2006 U.S. Dist. LEXIS 37465 at pp. 22-24 (D.Kan. 2006).

38. Electronic Discovery Reference Model, "Analysis – Pitfalls to Avoid", available at www.edrm.net/wiki/index.php/Analysis_-_Pitfalls_to_Avoid.

39. Electronic Discovery Reference Model, "Analysis – Updating Assessment as Issues Evolve", available at www.edrm.net/wiki/index.php/Analysis_-_Updating_Assessment_as_Issues_Evolve; Electronic Discovery Reference Model, "Analysis- Focusing Collection", available at www.edrm.net/wiki/index.php/Analysis_-_Focusing_Collection; Electronic Discovery Reference Model, "Processing – Searching", available at www.edrm.net/wiki/index.php/Processing_-_Searching; G.L. Paul & J.R. Baron, "Information Inflation: Can the Legal System Adapt?" (2007) 13 Rich. J.L. & Tech. 10 at paras. 50-55.

40. *The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production*, 2nd ed., June 2007 (Sedona, Arizona: The Sedona Conference, 2007) at p. 57, available at www.thesedonaconference.org.

41. M.R. Arkfeld, *Electronic Discovery and Evidence*, 2006-2007 ed. (Phoenix: Law Partner Publishing, 2006) at p. 6-25.

42. J.R. Baron, ed., "The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery" (2007) 8 The Sedona Conference Journal 189 at p. 214, available at www.thesedonaconference.org under "Publications".

43. Electronic Discovery Reference Model, "Review – Emerging Technologies", available at www.edrm.net/wiki/index.php/Review_-_Emerging_Technologies; Electronic Discovery Reference Model, "Review – Planning the Review", available at www.edrm.net/wiki/index.php/Review_-_Planning_the_ Review; R. Frederick, "Efficiency Matters – Using Topic Review Technology in Document Review" (2007) 8 Hearsay 5 at p. 6, available at http://www2.mnbar.org/sections/new-lawyers/Winter07.pdf; R. Nunez, "Automate Analysis During Review" (2006) 5 E-Discovery Advisor 20, available at http://e-discoveryadvisor.com/doc/18303; P. McLaughlin, "Manage Online Document Reviews for Electronic Discovery" (2006) 5 E-Discovery Advisor 16, available at http://e-discoveryadvisor.com/doc/18304; J.R. Baron, ed., "The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery" (2007) 8 The Sedona Conference Journal 189 at pp. 195, 215, available at www.thesedonaconference.org under "Publications".

44. Electronic Discovery Reference Model, "Review – Emerging Technologies", available at www.edrm.net/wiki/index.php/Review_-_Emerging_Technologies.

45. *Ibid.*; P. McLaughlin, "Manage Online Document Reviews for Electronic Discovery" (2006) 5 E-Discovery Advisor 16, available at http://e-discoveryadvisor.com/doc/18304.

46. Electronic Discovery Reference Model, "Review – Emerging Technologies", available at www.edrm.net/wiki/index.php/Review_-_Emerging_Technologies.

47. Electronic Discovery Reference Model, "Review – Planning the Review", available at www.edrm.net/wiki/index.php/Review_-_Planning_the_ Review; R. Frederick, "Efficiency Matters – Using Topic Review Technology in Document Review" (2007) 8 Hearsay 5 at pp. 6-7, available at http://www2.mnbar.org/sections/new-lawyers/Winter07.pdf.

48. Electronic Discovery Reference Model, "Review – Emerging Technologies", available at www.edrm.net/wiki/index.php/Review_-_Emerging_Technologies.

49. J.R. Baron, ed., "The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery" (2007) 8 The Sedona Conference Journal 189 at p. 212, available at www.thesedonaconference.org under "Publications".