

Evaluation of Intelligent Discovery Processes by Squire, Sanders & Dempsey

How the Firm Uses Equivio>Relevance to Reduce the Costs of Discovery

By Howard Nicols
and Stephen J. Goldstein

Due to the high volume of electronically stored information, document review and production is often the most expensive part of the discovery process. In an effort to lower the costs of litigation discovery, Squire Sanders invested in an assessment of next-generation intelligent discovery tools and processes. The purpose of this exercise was to identify and validate software and techniques that can defensibly reduce the expense of human review, while maintaining or improving quality. Having studied other available technologies, we decided to thoroughly evaluate the Equivio>Relevance system. Our results and general observations are as follows.

EVALUATION BACKGROUND

The test data was taken from a collection of documents related to the defense of a putative environmental class action related to a large manufacturing facility.

Howard Nicols is Global Managing Partner at Squire, Sanders & Dempsey L.L.P. and a core member of the firm's Value Partner Team, which focuses on improving business processes, knowledge management and project management to maximize client value. He leads the firm's Intelligent Discovery program, employing state-of-the-art technologies and streamlined processes to achieve and reduce electronic discovery costs. **Stephen J. Goldstein** is the firm's director of practice support. He is a member of the firm's E-Discovery & Data Management Team as well as a core member of the firm's Value Partner Team.

The documents used in the evaluation had been previously reviewed and prepared for production by our lawyers using traditional review methods. The case settled just prior to the actual production of the documents. Accordingly, there was a known end result to measure against.

The goal of this exercise was two-fold:

1. Test the ability of Equivio>Relevance software to identify known relevant documents within the evaluation review set; and
2. Evaluate the relative efficiency of Equivio>Relevance versus review by our lawyers.

The original review project consisted of approximately 200,000 documents collected from the client (50/50 mix of hard copy and electronic data). Using all available best practices to efficiently review the data, a team of nine lawyers took four months (1,250 hours) to screen and identify documents responsive to the opposing party's requests. While this effort was considered efficient at an overall average of 160 documents reviewed per billable hour, it still consumed nearly half a million dollars in billable time.

EVALUATION DATA

The evaluation focused on documents responsive to two similar requests for production.

1. 44,581 documents were loaded to the Equivio>Relevance system; and
2. The documents loaded had known results from previous human-review efforts.

Approximately 15% of the documents were known to be relevant and responsive to two of the 12 original requests for production. The remaining documents were generally nonresponsive, or responsive to other production requests beyond the scope of the evaluation.

EQUIVIO TRAINING

The Equivio>Relevance process works as follows:

- Documents are loaded into the Equivio>Relevance database.
- A senior, knowledgeable lawyer/ reviewer trains the system by coding a sample of 40 documents as relevant or nonrelevant.
- The Equivio>Relevance system analyzes the coding of the first sample and draws a second sample for coding.
- The process continues until the system concludes it has been fully trained (typically 45 passes of 40 documents for a single issue).
- Once trained, the system scores the predicted relevancy of all the documents in the system.
- The scoring takes place automatically using advanced algorithms and statistical methods described later.

Following the process above, a lawyer with advanced knowledge of the case reviewed 1,960 documents (49 sample sets) until the system stabilized. The training process took approximately 10 hours of dedicated review time. At the conclusion of training, the system scored the predicted relevancy of all 44,581 documents on a scale from 0 to 100 (with a higher number indicating a greater likelihood of relevancy). The scoring process took three minutes to complete.

EQUIVIO RESULTS

Once the documents have been scored, the system recommends the optimal F-measure, or cutoff, score, which is the highest balance between recall and accuracy. The F-measure can be used to decide which documents to consider nonrelevant (those ranked below the F-measure point). In the example in

Figure 1 on page 3, Equivio>Relevance determined the evaluation project had a cutoff score of 14. In this scenario, 74% of the documents were deemed to have a score too low to be considered relevant, or responsive to the request for production (*i.e.*, 74% of the documents could potentially be auto-culled without attorney review). In other words, the review could focus on the top-scoring 26% of the collection (documents with scores above 14), which, in this case, contain 94% of the relevant documents in the collection. **NOTE:** *The cutoff score can be adjusted to be more or less inclusive of the total data set in the judgment of the review project leader. Nonrelevant documents can also be sampled as a quality control measure.*

The validity of the Equivio> Relevance scoring is measured via a process known as discrepancy analysis. Typically, a sample set of a few thousand documents would be created and reviewed by lawyers. The results of the human-reviewed set would then be compared to the results of the Equivio>Relevance system in order to judge the accuracy of the results. The evaluation project had the benefit of drawing from a known production set previously screened by our lawyers.

The discrepancy set consisted of 1,398 randomly selected documents previously designated as responsive to our target issue. Additionally, 2,100 nonresponsive documents were randomly collected for the analysis. The results were compelling (see Figure 2 on page 3).

As shown in Figure 2, the Equivio>Relevance system agreed with the human review on 1,274 of 1,398 documents (91%). Likewise, the Equivio>Relevance system and the human-review team concurred 1,771 of 2,100 times (84%) on nonrelevant designations. Combined, the Equivio>Relevance system agreed with the previous human review on 3,045 of 3,498 documents (87%).

The next step in the process is to refer to the times when the Equivio>Relevance system does not match the human-reviewed set. From the 453 documents where Equivio>Relevance did not agree with the human review, we drew a sample of 100 documents to be analyzed by another lawyer with deep knowledge of

Figure 1

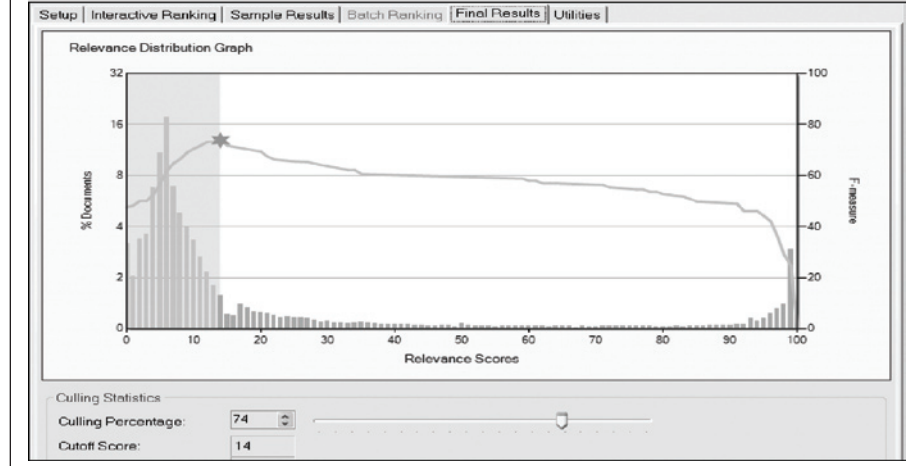


Figure 2

Figure 2 is a screenshot of a software interface showing a 'Discrepancy Matrix' table. The table has columns for 'Relevant (X)', 'Non-Relevant (Y)', and 'Total'. The rows are categorized by 'Reviewer' into 'Relevant (A)', 'Non-Relevant (B)', and 'Total'.

		Equivio>Relevance		
		Relevant (X)	Non-Relevant (Y)	Total
Reviewer	Relevant (A)	1274	124	1398
	Non-Relevant (B)	329	1771	2100
Total		1603	1895	3498

the case (a so-called “super-reviewer”). Again, the results were forceful.

As shown in Figure 3 below, of the 50 documents designated relevant by the human-review team and not relevant by Equivio>Relevance, only 18 were deemed actually relevant by the super-reviewer. Of the 50 documents judged not relevant by the human-review team and relevant by Equivio>Relevance, 21 were actually relevant. Based on this statistical analysis, the Equivio>Relevance system did as well as the human-review team (92% to 94% accuracy).

TRANSPARENCY

While there are a number of proprietary technologies at work behind the scenes, the Equivio>Relevance system is not a black box. The reasoning behind the learning and document scoring can be demonstrated via two critical tools.

1. Keyword generation. The Equivio>Relevance system automatically prepares a list of words/phrases contributing to, and against, responsiveness. Each term receives a positive or negative score that is

factored into the ultimate document ranking. Our evaluation suggests that these words and phrases can reveal knowledge about the case that would not otherwise be known or used in the early stages of a matter. **NOTE:** *This list could potentially be used in preparation for keyword negotiations with opposing parties.*

2. Transparency reports. Any document can be displayed with colors indicating the terms that contributed to, or against, relevancy. This tool is useful to assist in explaining the scoring methodology or in a potential response to challenges from opposing parties.

PROCESS

The quality and defensibility of an Equivio>Relevance review depends upon a well-defined implementation process. The following best practices and considerations should be incorporated into the review project.

TRAINING

The Equivio>Relevance training is only as good as the input of the skilled practi-

tioner who codes the document samples. The individual doing the training should have a high degree of knowledge of the case and code documents with precision and consistency.

ISSUE DEVELOPMENT

Defining the relevancy issues should be approached with care. If the issue is too broad or too narrow, the quality of results may be adversely impacted.

SAMPLING

Additional document sampling is desirable. A collection of random documents should be reviewed to create the required discrepancy analysis set (typically 3,000 to 5,000 documents). A second sample set can be used to test auto-culling decisions. In other words, if the Equivio>Relevance scores are leveraged to designate documents that scored below a cutoff as not relevant material, additional sampling and review can further validate the results.

DISCREPANCY ANALYSIS

In most cases, the built-in discrepancy analysis process is required to validate the Equivio>Relevance process and to bolster the defensibility of the review process.

CLIENT BUY-IN

The use of Equivio>Relevance affords a number of options. Clients should be well informed in choosing between these options. Just as with other review methods, clients should be advised and be given the opportunity for input regarding how the Equivio>Relevance tool will be used in a particular review project.

LEVERAGING EQUIVIO

There are a number of potential ways to cut review costs by leveraging Equivio>Relevance scoring:

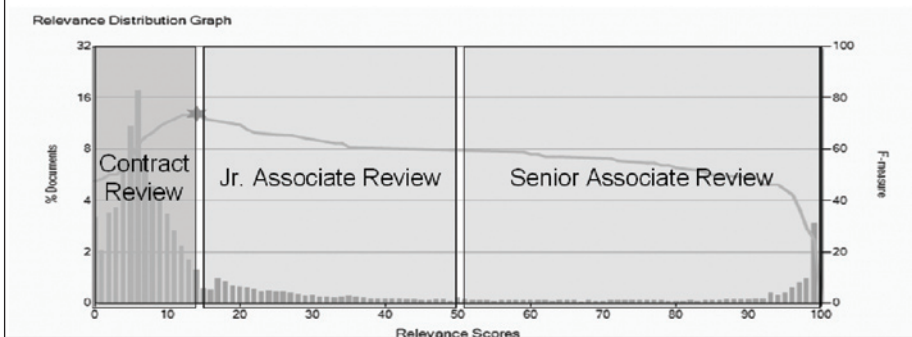
1. **Prioritize the review.** The documents can be sorted and reviewed from the most relevant to those not likely to be relevant. Reviewing in this fashion will allow the team to see the most relevant documents first and decide when to stop reviewing (*i.e.*, after a point where there no longer appears to be relevant or significant material justifying continued human review). Further, because the most significant documents will be grouped together

Figure 3

Expert Verification		Verification	
		Sample Size	Actual Relevant
Reviewer Relevant / E>R Non-Relevant (AY)		50	18
Reviewer Non-Relevant / E>R Relevant (BX)		50	21

Results	Recall	Precision	F-Measure	Accuracy
Equivio>Relevance	94.1%	88.1%	92.8%	92.0%
Reviewer	89.8%	96.4%	91.0%	94.2%

Figure 4



and seen first, the team will learn the case faster.

2. **Early case assessment.** Similar to the earlier example, the top X percent of documents can be quickly put into the hands of the decision makers to understand the client's position, which allows for early fight-or-flee determinations.
3. **Staffing.** As shown in Figure 4 below, different levels of staffing can be used for different score ranges. Senior law firm resources may look at the high scoring documents, while paralegal or contract staffing may be used to review documents with lower scores.
4. **Bulk-culling.** Using statistically valid approaches and proper quality control processes, Equivio>Relevance scores can be used to automatically cull a percentage of the document population.

CONCLUSION

Our evaluation of Equivio> Relevance suggests this technology can be used in a variety of ways to cut costs, improve accuracy and speed up the review of documents. It can also be used as a powerful early case assessment tool. At a minimum,

Equivio>Relevance offers the ability to augment the value of a typical discovery database. In other appropriate cases, it can be used (as part of the processes described earlier) to cull nonrelevant material significantly faster and more economically than traditional review methods.