
By Katie Winseck

Ms. Katie Winseck, Esq., is a solution architect for Kroll Ontrack, a global e-discovery technology and services provider (http://www.krollontrack.com).

The modern approach to reviewing and producing the deluge of discoverable electronically stored information, better known as e-discovery, has been significantly impacted by data proliferation, rising judicial intolerance of discovery blunders, and the increase in sanctions for mismanaging ESI. Companies are becoming increasingly familiar with e-discovery best practices and are now seeking innovative ways to minimize litigation costs while ensuring thorough and accurate review of documents produced in response to discovery and regulatory requests.

For the past decade, a linear, manual review process, with attorney eyes on every document, was hailed as the “gold standard” of document review methodologies. With increased reliance on email and other forms of ESI to conduct business, manual review was no longer a viable option and the use of keyword searching to narrow the universe of documents to be reviewed was born. Although technology has played a role in keyword searching, the “eyes on every document” approach remained intact, and litigants embraced keyword searching as a de facto extension of the gold standard. However, as the drawbacks of this approach become increasingly apparent, it appears that this rubric is quickly unraveling.

What is Technology-Assisted Review or Intelligent Review Technology?

In recent years, reliance on manual keyword searching has proven inefficient and in many cases, cost prohibitive. Many are looking to cutting-edge “smart” technology that offers a way to cut the cost of document review, which can account for 50 to 90 percent of discovery costs.

In light of these staggering numbers, attention is being paid to evaluating and leveraging technology-assisted review, or more specifically one form of the technology—Intelligent Review Technology (IRT), which includes Automated Workflow, Intelligent Prioritization, and Intelligent Categorization components.

Automated Workflow

Automated Workflow involves carefully mapping out each step of the document review process and defining which documents will be assigned to a particular group of reviewers prior to the commencement of the review. Automated Workflow allows parties to systematically identify groups of reviewers who are best suited to review certain documents and best leverage the strengths of those on the review team.

For example, a group of reviewers particularly strong in the area of privilege may be assigned to review all documents that have hit on search terms indicating a document is potentially privileged. Besides boosting the overall strength of a review team as a whole, use of Automated Workflow can result in reducing the cost of review management by up to one-third because reviewers “check-out” their document review assignments, rather than having them manually assigned to them by a paralegal or project manager.

Intelligent Prioritization
Intelligent Prioritization works in conjunction with Automated Workflow and is simply a way of organizing documents based on binary (e.g. “Responsive” and “Non-Responsive”) human decisions. Documents identified as potentially responsive will be elevated to the top of the review pile while documents that are potentially non-responsive will be pushed to the bottom of the pile, enabling parties to get to the most relevant, important documents much earlier in the review.

**Intelligent Categorization**

Last, but certainly not least, Intelligent Categorization is a software that conducts an automated and statistical analysis of attorney document coding decisions in relation to document content to suggest document categories based on the probability that document content will fit into a category that has been identified by human logic and applied based on attorney decisions.

In comparison to keyword searches, IRT is notable for yielding a more complete and accurate set of relevant documents while minimizing time and resources spent on review. The three components of IRT may be used together or a la carte, depending on the particular needs and issues of the case.

**Benefits of IRT**

One major benefit of IRT lies in statistical sampling, which a party uses to measure, monitor, control, and correct potential errors in categorization. Through frequent and detailed sampling, a party can effectively illustrate a high level of quality control throughout the review process while sampling results are supported with real time metrics.

Another benefit of IRT is that it might also be used in conjunction with reviewing documents from opposing counsel, to which parties often apply keywords to identify documents that can be used for purposes of deposition preparation and exhibits for trial.

**Continuing Reservations**

Despite the effectiveness of this and other similar emergent technologies, there are lingering concerns within the legal community regarding the efficacy of technology-assisted review platforms. Common concerns lie in misconceptions of how the technology works in practice, the lack of a judicial court opinion endorsing any such technology, and uncertainty regarding its defensibility.

Careful analysis of judicial treatment of keyword searches, however, reveals that courts have primarily focused on the reasonableness and defensibility of the process undertaken. The standard for evaluating a party’s use of keyword searching should arguably be applied when analyzing a party’s use of an automated review technology.

Accordingly, parties should adhere to best practices and discovery principles to ensure a thoroughly defensible e-discovery process regardless of the technology used.

**Keyword Searching in the Courts: The True ‘Gold Standard?’**

Although many consider keyword searching the "gold standard" for culling the size of a review set, this notion did not originate with the courts. In fact, judicial opinions have been highly critical of how parties have conducted keyword searches and have further noted that their role as triers of fact does not extend to endorsing a particular service provider or culling method.

For example, in *United States v. O'Keefe*, Magistrate Judge John F. Facciola noted that measuring the effectiveness of a particular method was a complicated question, and “for lawyers and judges to dare opine that a certain search term or terms would be more likely to produce information than the terms that were used is to truly go where angels fear to tread.”

He further stated, “This topic is clearly beyond the ken of a layman and requires that any such conclusion be based on evidence that, for example, meets the criteria of Rule 702 of the Federal Rules of Evidence.” 1 Facciola re-stated this sentiment in *Equity Analytics LLC v. Lundin*, noting that identifying the effectiveness of a particular search methodology was “beyond the ken of a layperson or lawyer.” 2

---


**Focus on ‘Reasonableness’**

Moreover, rather than outlining a bright-line standard, most decisions discussing the use of keyword searches and other technology in litigation have focused on whether the process undertaken was reasonable. Magistrate Judge Paul Grimm, a leading voice and authority on e-discovery issues, stated that "When Parties decided to use a particular ESI search and retrieval technology, they need to be aware of the literature describing the strengths and weaknesses of various methodologies ... and select the one most appropriate for its intended task.” 3
Grimm’s opinion, and others like it, suggest that the courts’ primary concern with regard to the defensibility of a particular technology lies in the amount of diligence and proactive planning put into leveraging technology, rather than the technology itself.

**Need for Cooperation**

In addition to careful consideration of how search technology has been used, courts have also assigned a great deal of importance to cooperation between the parties and proof of a particular methodology’s efficiency. In *William A. Gross Construction Associates Inc. v. American Manufacturers Mutual Insurance Co.*, Magistrate Judge Andrew Peck—probably the biggest judicial advocate for use of technology to assist in document review—identified the importance of “carefully craft[ing] appropriate keywords, with input from ESI custodians” and noted that “the proposed methodology must be quality control tested to ensure accuracy and elimination of ‘false positives.’”

Endorsement of Specific Technology Unlikely

Taking these opinions as guidance, it is unlikely that courts will set forth a blanket endorsement with regard to the use of any single provider’s specific technology-assisted review platform, as none of these opinions explicitly condones any review technology. Instead, each opinion demonstrates that the courts lack sufficient knowledge in the context of keyword searching—an approach that has evolved as a best practice standard over the last 10 years—and have instead focused on determining whether the party implementing the process did so defensibly and adhered to discovery best practices and principles.

**IRT as the Way of the Future**

In addition to concerns about judicial approval, many express concern that technology-assisted review is devoid of human control and eliminates the “eyes on every document” approach. This concern arises from the misconception that technology-assisted review platforms are “black box” technologies that eliminate the element of human review. In reality, some level of human interaction is at the core of most technology-assisted review platforms.

**Distinguishing Technologies**

Indeed, there is undoubtedly much confusion of current “smart” versions of the technology, such as IRT, with the much more widely-known concept clustering or topic grouping technologies that were launched over five years ago. These older, less “smart” technologies are used by parties to eliminate large swaths of documents from review altogether with little human involvement in the document clustering. Newer, “smarter” technologies, such as IRT, require an extensive amount of human intervention to train the system to learn from human decisions.

**Precision and Recall**

Additionally, case studies suggest that IRT is extremely effective at producing a high number of relevant documents (recall), while also producing a highly accurate fraction of irrelevant documents (precision)—especially in comparison to using search terms alone. A 1985 study by David Blair and M.E. Maron found that experienced attorneys and paralegals that used keywords to search over 40,000 documents had an average recall of 20 percent. In other words, attorneys only found 20 percent of the relevant documents through keyword searching.

Ongoing research by the Text Retrieval Conference (TREC) echoes earlier findings in the Blair & Maron study, concluding that Boolean keyword searches found only 24 percent of the total number relevant documents. Further, a 2010 TREC study found that 76 percent of relevant documents were retrieved when automated review technology was used.

In addition to high recall, a study by the Discovery Institute found that using technology-assisted review reduced costs by 45 percent or more.
Sedona's View

In addition to data supporting the overall efficiencies and increased search accuracy achieved through use of automated review technologies, such as IRT, many e-discovery authorities now support the use of this technology. Recently, the Sedona Conference issued an open endorsement, stating that “reliance solely on a manual search process for the purpose of finding responsive documents may be unfeasible or unwarranted.” 8 It was further noted that automated search methods can be viewed as “reasonable, valuable, and even necessary” to reduce costs and “safely and substantially reduce the amount of ESI that must be reviewed by a human.” Applying this commentary to the criteria used by the judiciary in law analyzing keyword searching, it is clear that the standard of review is one of reasonableness in connection with how the technology is used.

Judge Peck as Pioneer

As it currently stands, the use of automated review technologies is being heavily scrutinized by members of the bench. For example, in 2011, Magistrate Judge Andrew J. Peck opined significantly on this matter. At both the Carmel Valley e-discovery retreat and in writing, Peck openly questioned why lawyers were so insistent on waiting for the bench to “bless” smart review platforms. 9 His analysis identifies the inefficiencies of keyword searching, and explicitly points out that the judiciary never endorsed keyword searching—or any other review technology.

Instead, Peck echoes sentiments of existing case law, noting that, regardless of the technology used, he would examine the process undertaken by the party and evaluate whether it involved adequate quality control, sampling, and validation of recall and precision.

On Feb. 8, 2012, in a status conference transcript in Da Silva Moore v. Publicis Groupe, Peck addressed the use of this technology by assisting parties in implementing a protocol they developed to leverage one form of technology-assisted review: predictive coding. 10

In a formal written opinion by Judge Peck issued just a few days later on February 24, 2012, he states:

What the Bar should take away from this Opinion is that computer-assisted review is an available tool and should be seriously considered for use in large-data-volume cases where it may save the producing party (or both parties) significant amounts of legal fees in document review. Counsel no longer have to worry about being the ‘first’ or ‘guinea pig’ for judicial acceptance of computer-assisted review. 11

For parties considering use of technology-assisted review platforms, yet still reluctant to try them, such commentary from Judge Peck and other members of the bench should be heavily considered and continuously monitored.

Demonstrating Defensibility, Regardless of Method

Ultimately, questions regarding defensibility rely less on the technology used and more on the steps taken to implement it. Should a challenge arise, a party should be prepared to provide a detailed explanation that outlines each step of the processes undertaken, as well as the rationale for each step.

Make the Process Iterative

Additionally, there are specific measures parties can take when implementing these technologies. With increasing amounts of empirical data suggesting the inefficiencies of keyword searching, parties should employ an iterative process that assesses and continually hones the syntax and search criteria used to maximize recall and precision.

Parties should also sample non-produced documents early and throughout the review to demonstrate that adequate quality control was built into the process.
**Remember the Human Factor**

There are also a handful of options to enhance the defensibility of a process that uses IRT, including seeking agreement from opposing counsel, consulting with an IRT professional services expert, documenting the process followed, following the process, and making changes when the need arises. This methodology requires constant active human interaction and monitoring.

Finally, as judges are reluctant to favor parties who employ a “black box” technology without a full understanding of it, parties should be prepared to explain the technology to a judge or layman—either through an expert or individually.

**Conclusion**

Despite widespread concerns about using automated review technologies, savvy organizations and firms are quickly realizing that this technology is designed to aid a human document review—not replace legal analysis and advocacy. When properly leveraged, the cost savings are substantial and the process is defensible.

As stated by Judge Peck in *Da Silva Moore*, "It certainly works better than most of the alternatives, if not all of the alternatives. So the idea is not to make this perfect, it’s not going to be perfect. The idea is to make it significantly better than the alternative without nearly as much cost."

Prospectively, the key will lie not in the technology used, but rather the steps taken to marry it with a robust process with high recall, precision, and sufficient quality control.