

Continuous Active Learning for Litigation Discovery

Continuous Active Learning for Litigation Discovery: Enhancing Analysis and Revealing Essential Insights

By: Cody Gavalier

There is no denying that the technological revolution has had a substantial impact on virtually every major industry, including the legal sector. While many technologies have influenced the legal landscape, one, in particular, is forever altering the litigation discovery process: Continuous Active Learning, or CAL.

CAL has had the most significant impact on the document review process, so much so that it has become a necessity in the current legal ecosystem.

What Is CAL?

Artificial Intelligence is increasingly being used to analyze documents and predict whether they're relevant to any specified criteria. Continuous Active Learning, or CAL, is the most effective and popular method of training Artificial Intelligence for use in eDiscovery. As the name suggests, the model learns continuously, i.e., it updates its predictions regularly with new coding decisions from human reviewers. While there are no other technologies quite like CAL, the processes it uses, and the outputs it produces, can perhaps best be compared to the popular music application "Pandora."

When using Pandora, the application closely monitors the artists and genres of music that you listen to most frequently while providing a "thumbs up" or "thumbs down" to specific songs. Over time, Pandora becomes exceedingly efficient at presenting you with more of the songs and artists that you like while screening out music that you have responded negatively to in the past. Similarly, CAL learns what is relevant from reviewer feedback, makes suggestions based on that feedback, and continuously updates its predictions until it reliably provides reviewers with relevant documents. It is also active in its own learning process, selecting documents for human review that it believes will be most helpful in improving its predictions or getting the most relevant documents in front of the review team fastest.

This straightforward tool can be used in an astounding variety of innovative and valuable ways, but its use is still not widely adopted across the industry. In fact, [in a recent survey](#), 36% of practitioners were not using CAL at all. Instead of shying away, there are several use cases where deploying CAL solutions can streamline the document review process. This offers a multitude of benefits to both practitioners and their clients. This technology helps attorneys honor their fiduciary duty to their clients by verifying the thoroughness of the discovery process and providing faster access to insights and evidence.

Use Cases for CAL

CAL solutions are not meant to replace the human element in document review or litigation discovery. In actuality, CAL solutions enhance the effectiveness of the human component by building a predictive model to rank documents based on relevance. These rankings can then be used to augment or supplement the review process.

CAL for Quality Control

One common myth is that CAL requires an all-or-nothing approach—that human review is automatically eliminated once a litigation team buys into CAL. The reality is that CAL augments the human review for purposes like quality control, providing an AI-generated opinion about a document’s coding to double-check against human reviewer decisions.

In doing so, the model can help uncover presumptively incorrect documents beyond that capability of keywords and targeted searching.

Firms and legal departments should look to deploy CAL for quality control (QC) as a low-risk method of adoption since attorney eyes can still be placed on all documents. With near-immediate benefits and minimal setup to achieve these results, this approach can lead to as much as a 30% decrease in total QC hours, which lowers overall review cost without eliminating the human element from the review process.

Avoiding Irrelevant Documents

Parties seeking to control litigation spend can look to CAL to drive significant cost savings by reducing the need for human review or augmenting workflows to incorporate global review teams.

In a recent case involving more than 4 million documents, CAL use immediately saved roughly 60,000 review hours (the equivalent of 250 months of full-time work) and an estimated \$3.5 million in upfront fees. Subject matter experts trained the CAL model for responsiveness, and immediately more than 900,000 documents were excluded from review entirely.

While elimination of first pass review is most commonly considered, CAL opens the door to risk-balanced work allocation using global teams. Hesitant parties may elect for human review of all documents; however, CAL can help segregate presumptively non-responsive documents for human review in low-cost geographies. This can be a very effective method of reducing overall cost while introducing in-house counsel and case teams to the benefits of CAL and offshore review.

Speed to Intelligence

CAL can also help with litigation preparation by prioritizing presumptively relevant documents for human review and doing so before review fees begin to pile up. Documents that can make or break cases are served up quickly through prioritization. CAL accelerates the process but doesn’t eliminate the use of human review because it simply recognizes which documents need human eyes first.

In this use case, CAL saved both time and money by eliminating the need for the review team to search through unnecessary information, like bulk calendar items. Prioritized review is similar to getting a second opinion, with CAL as the first review option.

CAL can also leverage existing work product to provide rapid insight to new case data. For example, key or “hot” documents identified during review can be used to train a CAL model that is then applied to inbound opposing party productions to identify records relating to the same key concepts.

The Benefits of CAL

Continuous active learning review tools can be quickly calibrated to prioritization and work allocation.

Depending on use, CAL can provide significant cost savings to clients and enhance the overall accuracy of review. The specific benefits of CAL review tools include:

CAL Learns in as Little as Five

Perhaps the most notable advantage of CAL technology is its ability to learn. Instead of the protracted learning process during manual, linear review, CAL models can be trained quickly and iteratively throughout the review process. A human reviewer must manually start the process to provide the model feedback, however, the algorithm can begin predicting ranks after review of only five relevant and irrelevant documents. This initial sample helps the algorithm establish a baseline, which is then used to queue up additional relevant documents. As the reviewer progresses through the document assessment process, the software will make more and more accurate suggestions.

CAL Facilitates First-Pass Elimination and Enhances Efficiency

As shown in the previous real-world use case, another significant advantage of using CAL is that it can facilitate the elimination of first-pass review. Traditionally, first-pass reviews constitute a significant source of wasted time and resources. CAL eliminates this barrier to productivity by actively learning and prioritizing relevant documents allowing smaller, specialized teams of highly qualified human reviewers to focus on privilege, hot documents and other areas of vital importance to the case.

In addition, CAL allows attorneys to assess relevant documents more rapidly, which speeds and facilitates the entire case preparation process. This streamlining leaves them with more time to participate in dynamic tasks, such as conducting investigations, planning case strategies, or interacting with clients.

CAL Can Look Beyond Keywords

Finally, CAL can look beyond keywords to queue up documents relevant to the review process. A recent case study examined the impact of CAL's review and queue capabilities.

Though search terms provide clarity for the review process, you don't want an algorithm to ignore a "lunch meeting" because the search term was looking for "dinner." In this case, CAL can eliminate obviously unnecessary documents, while also serving up "hot" documents in which relevancy occurs. In a case study of 80,000 documents, CAL served up a little more than 4,000 that were "hot" enough to need review.

In the meantime, this search saved an estimated 2,000 review hours and approximately \$65,000.

Overcoming Challenges of CAL Deployment

Technology assisted review including CAL is well accepted by the courts since at least 2015, when U.S. Magistrate Judge Andrew J. Peck noted in *Rio Tinto PLC v. Vale S.A.*, 306 F.R.D. 125 (S.D.N.Y. 2015), that “the case law has developed to the point that it is now black letter law that where the producing party wants to utilize TAR for document review, courts will permit it.” Despite the courts’ acceptance and significant benefits associated with using CAL for litigation discovery and QC purposes, it remains that 36% of legal teams that haven’t used any of these technologies in their practices. Successfully implementing any new technology involves overcoming a set of internal and external challenges, such as:

A Lack of Organizational Buy-In

One of the most notable barriers to deploying CAL technologies is a widespread lack of buy-in from outside counsel and litigation teams. Qualms over an all-or-nothing approach coupled with the elimination of the human element remain, however, these fears are unfounded. A well-trained review team coupled with the technical and workflow expertise to deploy this powerful technology has proven itself across litigation types and industries. A lawyer’s duty to communicate under ABA Model Rule 1.4 requires that attorneys “reasonably consult with the client about the means by which the client’s objectives are to be accomplished.” Fortunately, fulfilling this requirement with respect to CAL is usually as simple as presenting stakeholders and decision-makers with relevant case study data.

Upon observing the substantial time and cost-saving benefits of CAL solutions, the overwhelming majority of stakeholders will be open to at least testing the technology for review purposes.

The Worry of “Dirty Data”

If CAL technology has any real shortcoming, it is the “garbage in/garbage out” phenomenon that plagues any active learning solution.

CAL technology is extremely resilient, however, the efficacy of CAL solutions are reliant on the input of the reviewer. CAL is most effective when the reviewer is a subject matter expert that exhibits consistency during the learning process. CAL’s resilience can eventually overcome an initial lack of subject matter expertise as reviewers gain familiarity with a matter, but the time to the model’s peak efficiency will be affected. Therefore, all team members must be trained on the technology prior to deployment. Organizational leadership should also establish document review best practices to ensure that all items are assessed using standardized guidelines.

Using CAL Is Necessary for Litigation

Due to the sheer complexity and volume of modern data sets, using CAL has become a necessity for in-house attorneys and private firms. The courts have frequently upheld the use of continuous active learning over simple keyword searches because of its capabilities.

CAL provides the most pragmatic solution for increasing organizational time and fiscal efficiency, enhancing the accuracy of the document review practices, and minimizing the need to use time-consuming manual processes.

For organizations searching for a way to enhance litigation discovery and document review practices, CAL is the solution. ABA Model Rule 1.1 requires attorneys to bear a duty to provide competent representation to their clients. This includes the duty to stay “abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology.” (ABA Model Rule 1.1, cmt. 8.) Given the current acceptance in courts and wide range of potential use cases, the time may soon come when failing to consider CAL will be viewed as a failure of the practitioner’s ethical obligations.



want you first.
We second that.

We don't have a one size fits all approach, because clients are like cases. No two are alike.

Learn more at
Unitedlex.com/litigation

