



## CASE STUDY

# High-Efficiency Review Enables Client to Meet Short Deadlines; Significantly Reduces Costs

Catalyst Managed Review + TAR 2.0 Finds 94% of the Relevant Documents

“ *The Catalyst review team achieved a remarkable review efficiency and looked at very few nonresponsive documents.* ”

### Client Snapshot: Global Corporation

- Federal investigation into alleged price fixing
- 2 million documents to review in less than four months
- Highly responsive collection with changing review criteria
- Client saved \$400,000 by using Catalyst's technology and review team

## Background

Our client, a global corporation, was facing an investigation by a government agency into alleged price fixing. The regulators believed they would find the evidence in the documents and issued a broad request—with just four months to get the job done.

This wasn't a case of finding a needle in a haystack. Rather, a wide range of documents were responsive. A sample of the initial collection suggested that as many as 45% of the documents would be responsive. One option was to produce everything but the client had a significant amount of confidential and proprietary information that it did not want to be inadvertently exposed. Thus, they needed to produce responsive documents, but only responsive documents.

How would the company get through the documents, finding the responsive (and only the responsive) documents to meet the production deadlines, including one within the first two weeks of the review—compounded by numerous unforeseen changes in review criteria and last-minute rolling collection uploads that lay ahead?

## Making Review Efficient

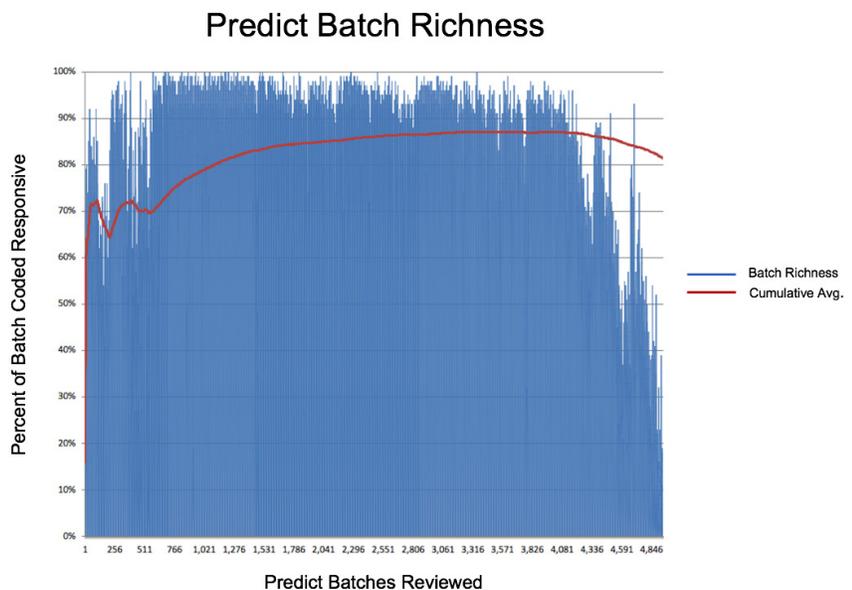
Our goal was simple: Use Insight Predict, our technology-assisted review (TAR 2.0) continuous active learning (CAL) engine, to find the relevant documents as quickly and efficiently as possible. We matched our engine with the Catalyst review team, experienced Predict document review experts who are versed in getting the most out of our software.

We used CAL from the beginning, without the need for a senior lawyer subject matter expert (SME) to review thousands of documents to train the system. (TAR 1.0 systems require significant up-front training.) Rather, the team got started right away using the responsive documents already identified for initial training, with no time to waste.

The measure of a predictive review is how quickly the algorithm can surface relevant documents. Like a bloodhound born to track, Insight Predict picked up the trail almost immediately.

Here is a chart showing the percentage of relevant documents Predict found on a batch by batch basis.

There were almost 5,000 batches in this review. Each blue line represents the percentage of relevant documents in the batch. In the early stages, the number reached 80% to 90%. That meant the Catalyst



review team found 80 to 90 responsive documents out of 100 in their batch. It also meant that they saw only a few nonresponsive documents, which was our ultimate goal. Make the reviewers as efficient as possible, to keep review costs as low as possible.

## Review Efficiency

This is an important but seldom-discussed topic. How many nonresponsive documents does the reviewer see for each responsive one? As stated, our goal was to maximize the number of relevant documents our team reviewed while minimizing the nonresponsive ones.

When keywords are used to cull documents, reviewers typically have to look at as many as nine nonresponsive documents for each responsive one. That means they are wasting their time for about 90% of their review efforts. With Predict, our statistics show that the ratio is much narrower, about 2 to 1, which means the team finishes faster and bills less. We call this “review efficiency,” and it is an important ratio to consider when looking at e-discovery alternatives.

In this case, our team achieved a review efficiency of 1.33 to 1, which is pretty remarkable. That means that the review team looked at very few nonresponsive documents over the course of the project. Taking the math up two decimals, the average team member only had to look at 133 documents to find 100 responsive ones. This remarkable efficiency was achieved by the combination of the TAR 2.0 engine plus the expertise of our reviewers.

That’s why the chart shows batches that quickly approached 100% responsive. Interestingly, you will see a responsive rate dip early on in the project. That was because the team had to finish reviewing a few key custodians’ documents first and when they started running out of relevant documents, the numbers dipped. When we opened the review back up to the entire collection, the responsive rate jumped as well. Some batches were 100% responsive.

You can also see that batch richness dropped at the end. This is to be expected. With CAL, the goal is to keep reviewing until you stop seeing responsive documents. Once the responsive rate drops substantially (say to a tenth of the high-water mark), that is a signal that it is time to stop. In this case, our team kept reviewing batches to make sure they were nearing the end.



The average team member had to look at just **133 documents to find 100 responsive ones.**

At the end of the day, the Catalyst review team reviewed 600,000 documents.

## High-Quality Review Results in Cost-Savings

Toward the beginning of the review, the client's outside counsel conducted a quality control (QC) review of 20% of the documents passed to them by our review team. Seeing the exceptionally high quality of the review, they quickly reduced their QC review to just 2.5%. This included our review of 35,000 potentially privileged documents and drafting a 10,000 entry privilege logs. Then, through the course of the project, the law firm's QC review was reduced to just 2.5% due to our high-quality work. Over the course of the review, 15,970 documents were QC'd by outside counsel and only 848 were changed (95% accuracy) and almost all were changed during the first two weeks of the review as Catalyst used its QC workflow to ensure that our team's coding was aligned with that of outside counsel.

Here is a chart showing the cost-savings for the 600,000 documents we reviewed:

Outside Counsel QC Percentage	Number of Docs for Outside Counsel Review	Outside Counsel Review Rate	Outside Counsel Hours	Outside Counsel Billable Rate	Cost
Standard: 20%	120,000	100 documents/hour	1,200	\$400/hour	\$480,000
Catalyst Managed Review: 2.5%	16,000	100 documents/hour	160	\$400/hour	\$64,000

## Meeting Short Deadlines

Initially, our team needed to complete review of 45,431 documents to meet a two-week production deadline for the initial set of documents. We met the deadline with no problem, working through two consecutive weekends.

As new deadlines arose following the initial production, we doubled our team to more than 40 reviewers to meet deadlines.

## Changing Review Criteria

Then came a wrinkle in the review process, although it is not uncommon. At a couple of points along the way, team leaders refined their view of responsiveness. This is a natural process as you learn more about your documents and about your case. Many call it relevance drift. The simple fact is that you know more about your needs at the end of the process than at the beginning. It is one of the biggest weaknesses of the old TAR 1.0 process. If all your training is done at the beginning, how do you account for what gets learned as you go along?

Catalyst's TAR 2.0 algorithm is noise tolerant, which means that coder inconsistencies—and even changes in direction—do not adversely affect it. With CAL, every ranking starts fresh, with no memory of the previous one. That way, if you were to re-tag tens of thousands of documents, the next ranking would take it in stride. The same is true as the team refines its search objectives.

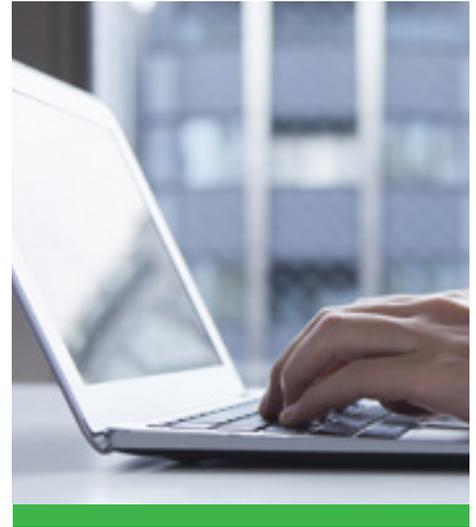
That happened several times in this case as understanding increased. Yet, our team successfully navigated the multiple major changes to the review protocol with no adverse impact on performance.

## New Data Uploads Three Days Before the Production Deadline

Finally, just three days before the final production deadline, the team was tasked with reviewing contact lists for the 15 most important custodians. We were able to review over 10,000 entries in the short turnaround.

## The Results

The team achieved 94% recall in this review, far greater than that required by the courts. We did so having reviewed only 30% of the collection, meeting our client's deadline with room to spare. Review costs were a lot lower than they would have been with linear review or keyword search.



**Learn More**  
[catalystsecure.com](https://catalystsecure.com)



**Follow Us**  
[@catalystsecure](https://twitter.com/catalystsecure)