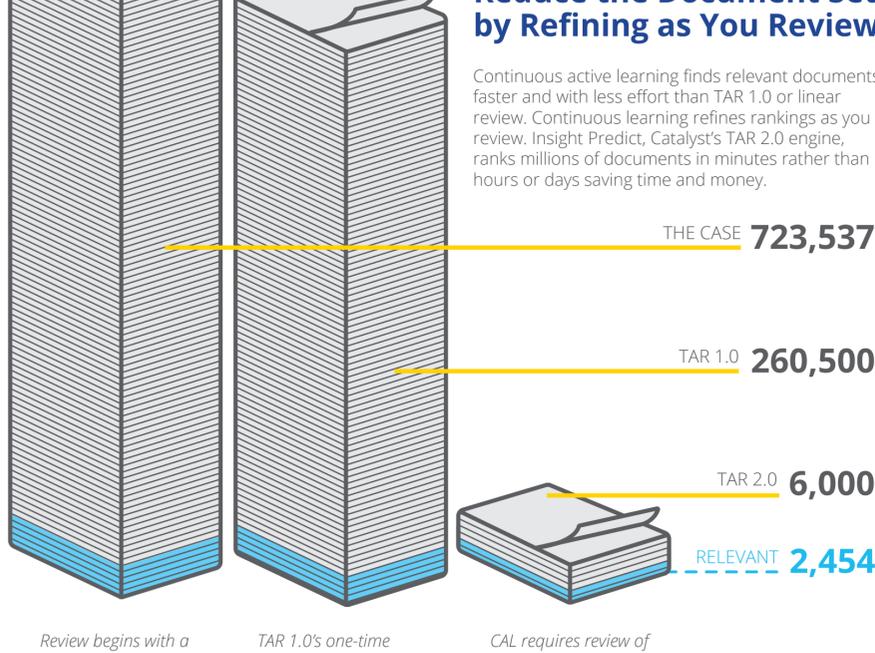


CUT THE COST OF DISCOVERY IN FIVE easy steps

1 USE CONTINUOUS ACTIVE LEARNING To Reduce Your Document Set

Reduce the Document Set by Refining as You Review

Continuous active learning finds relevant documents faster and with less effort than TAR 1.0 or linear review. Insight Predict, Catalyst's TAR 2.0 engine, ranks millions of documents in minutes rather than hours or days saving time and money.



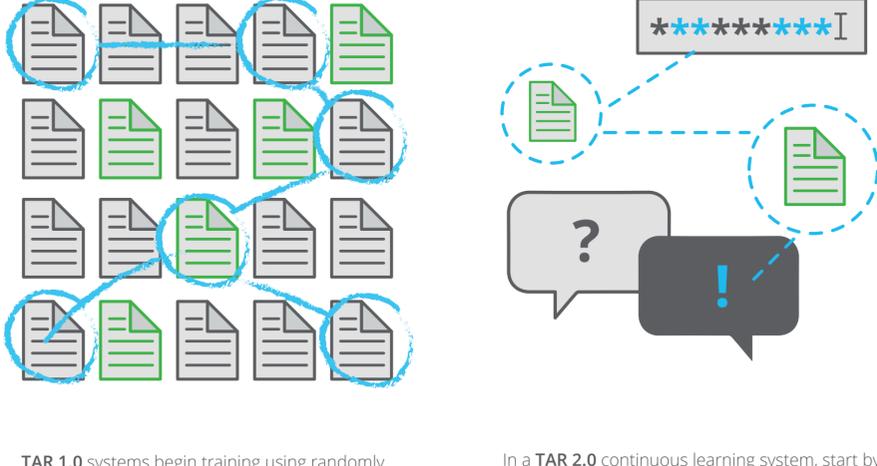
Review begins with a collection of 723,537 documents.

TAR 1.0's one-time training would reduce the collection, but still leave 260,500 to review.

CAL requires review of only 6,000 documents to find 75% of the 2,454 that are relevant.

2 STOP USING RANDOM DOCS For Training the TAR System

KICKSTART RANKINGS WITH RELEVANT DOCS



TAR 1.0 systems begin training using randomly selected documents. Creating a control set requires that you tag 500+ randomly selected documents. Then you train based on random documents of marginal relevance as the system tries to determine a "cutoff" point for review.

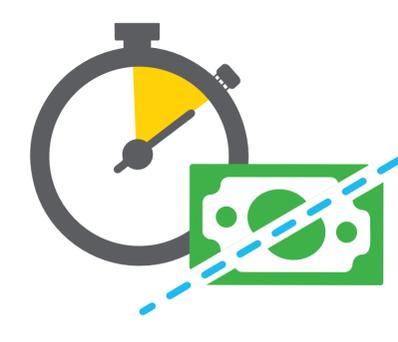
In a TAR 2.0 continuous learning system, start by finding as many relevant documents as you can, through keyword search, witness interviews, past productions or otherwise. Use those as training seeds to start the ranking process. The review team focuses on likely relevant documents from beginning to end, cutting review costs and time.

Independent research conclusively shows that this is the **least efficient way to run a review**—particularly when there aren't many relevant documents in the collection (low richness). Your trainer spends hours or days looking at marginal or irrelevant documents while the review team waits.

3 START WITH REVIEW TEAMS Not Senior Lawyers

TAR 1.0
Requires a subject matter expert—typically a senior lawyer—to review thousands of documents to build a control set, train the system and test results. Because the review team can't start until the senior lawyer finishes, this causes delays and added expense.

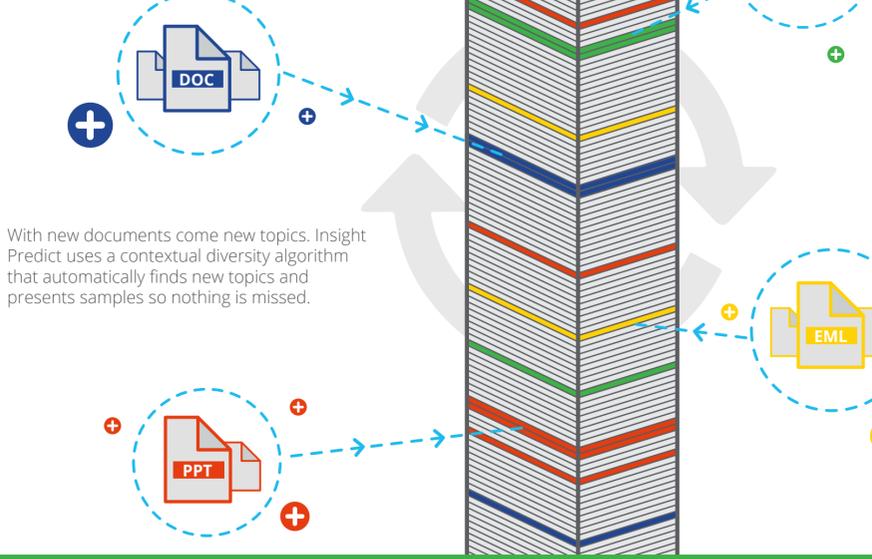
TAR 2.0
Continuous learning eliminates the need for training by a senior lawyer. It learns from the reviewers so the review can start immediately. As reviewers progress through documents, the system continuously uses their judgements to improve its rankings, dramatically reducing both cost and time.



4 ADD NEW DOCUMENTS At Any Point

ROLLING COLLECTIONS

Because Insight Predict doesn't use control sets, you can add more documents at any time and have them join in the ranking. With TAR 1.0 systems, you would have to create a new control set and repeat the training.



With new documents come new topics. Insight Predict uses a contextual diversity algorithm that automatically finds new topics and presents samples so nothing is missed.

5 AVOID CONTROL SETS With Continuous Active Learning

PROBLEMS WITH CONTROL SETS

REPRESENTATIVENESS

TAR 1.0 systems use control sets of 500 or more documents to represent larger collections. This is neither statistically accurate nor logically realistic. The problem gets worse when richness is low.

KNOWLEDGE

As review progresses, understanding of the case improves. Early relevance calls might be made differently later. Basing review on a control set tagged at the beginning can severely limit the algorithm, requiring more documents than necessary to be reviewed.

SUBOPTIMAL

Using control sets for a review cutoff is likely to yield poor results. Stability against a 500-document control set doesn't equate to stability against a much larger set. Either recall is poor or you end up reviewing more documents than necessary.

BONUS: USE CONTEXTUAL DIVERSITY And Find Relevant Documents Faster

KNOW WHAT YOU DON'T KNOW

Contextual diversity is a unique feature of **Insight Predict** that constantly explores data to find examples of content that reviewers haven't seen. With every ranking, the system presents a mix of relevant and unexplored documents, so reviewers are constantly considering new pockets of data.

As not miss key, but unexpected documents that you might not have found through keyword search or general investigation. Contextual diversity groups these unseen documents into clusters and includes representative samples in the review mix.