

Analytics & Predictive Coding

Tools that help speed and prioritize reviews, save money and improve production quality.

In order to keep up with the explosive growth of unstructured data, corporations and their outside counsel require tools that can filter through the noise of non-responsive documents, help streamline and prioritize documents for review, and ensure quality throughout the eDiscovery workflow. Advances in text classification technologies have not only made such tools an actualization, but also readily available. Consilio's unique approach integrates these tools into a high-caliber review platform, Global RPM, and *further* leverages them with the expert hands of a seasoned master project manager to serve any complex matter.



There are many advanced analytics capabilities available to users in the Consilio review platform, including Conceptual Search, Conceptual Clustering, Email Threading, Near Deduplication and Predictive Coding. Through leveraging our technology partner's best-in-class technologies and integrating them with our own,

Global RPM, Consilio makes these unique solutions available on a global basis so as to better serve our client's most complex multi-jurisdictional and multi-lingual data set eDiscovery needs.

Conceptual Search and Conceptual Clustering

By applying Latent Semantic Indexing algorithms to a corpus of data, Consilio is able to expose, through our review platform, the ability to search for concepts within a data set that goes beyond the limitations of syntax-dependent Boolean search queries. This enables clients to find documents that are conceptually relevant even if there are no keyword hits, and also gives clients the ability to allocate conceptually-related documents for review; thus, improving the efficiency of the review team. Furthermore, clients can explore the data corpus *visually* using Global RPM's innovative Concept Wheel which shows spatial relationships between concepts in the corpus.

Email Threading

With e-mail threading, documents that are part of a sequence of e-mails are related in the database so that clients can look at the contents of the entire chain to speed up the review cycle. In addition, clients can view the "most inclusive" thread in a conversation in order to quickly determine the responsiveness of the entire thread family. E-mail threading also allows for an additional QC layer to be incorporated into the workflow, ensuring similar coding across all documents in the thread.

The Right Tool for the Job

Every craftsman knows selecting the right tool for the job is the difference between a job well done and a frustrating learning experience. Consilio's seasoned project managers have the depth of experience and technical understanding to guide your eDiscovery project to use each of these tools as appropriate to create the most efficient, cost-effective, defensible workflow for your matter.

Conceptual Search: Best applied in a "find more like this" scenario, or for prioritizing document subsets for review or to help cluster similar documents for more efficient review.

Concept Clustering: Best for exploratory investigation of the corpus to understand "what do I have?", and for understanding spatial relationships between concepts in the corpus.

Email Threading: Useful in grouping threaded discussions for review, and ensuring coding decision continuity for threaded documents.

Near Duplicates: Good for culling the data set, batching near duplicates for more efficient review and QC of coding decisions.

Predictive Coding: Excellent for knocking out the non-responsive docs, prioritizing review resources and pre-production QC.

Near De-duplication

With near de-duplication technology, Consilio can identify documents that are highly similar to one another, but not perfect duplicates. Doing so affords the review team the option of batching near duplicates for improved review efficiency, and also affords review managers another option to QC near-duplicates to ensure coding continuity prior to production.

Predictive Coding

One of the most exciting trends in eDiscovery is the broad adoption of computer modeling to leverage a subject matter expert's coding logic across a very large corpus of documents, without experiencing reviewer inconsistency or fade.

Consilio has integrated predictive coding into our review tool, Global RPM - alongside other advanced analytics tools - so that clients can

enjoy the benefits of predictive coding workflows without the challenge of deciphering the complex language associated with predictive coding, or being quantitative statisticians themselves.

Consilio's predictive coding workflow is flexible, unlike the hard-coded and overly prescriptive workflows of many

other solutions on the market. Clients can enjoy predictive coding's benefits as it makes sense for their workflow - for early case assessment, for culling non-responsive documents, for prioritizing review, or for pre-production QC format. The guided, personalized approach offered by Consilio adjusts for each client's desired confidence intervals, low prevalence data sets, selection of non-random "key documents" to accelerate the training, rolling data collections or rolling productions, and concurrent modeling of multiple tags. The net result is a workflow that suits each client's resources, constraints and interests.

Consilio's predictive coding offering fully supports multi-lingual data sets. With proven support for Japanese, Chinese, Portuguese, Swedish, English and more, Consilio's predictive coding technology utilizes a highly sophisticated classifier combined with a Unicode-compliant data store to efficiently model documents of a different language without having to separate the data set by language, or create language-specific computer models.

The Right Tool for the Job

There are a lot of predictive coding tools on the market, so why select Consilio? Here are the top 4 reasons:

Reason #1: Unified Review Platform. Unlike others, Consilio's predictive coding capabilities are integrated into Global RPM - the industry's leading review platform. So alongside your search, assignments, foldering and redactions, you have the ability to access all advanced analytics and predictive coding capabilities.

Reason #2: Global Solution. For complex multi-jurisdictional and multi-lingual matters, clients have the benefit of a technology stack, hosting centers, and project management to support matters around the globe.

Reason #3: Flexibility. Our predictive coding solution is not hard-coded into a workflow that's rigid and overly prescriptive, so our expert project managers can help clients bring predictive coding to bear in a way that best serves the client's eDiscovery interests.

Reason #4: An Advantaged Technology. Consilio's predictive coding engine is self-tuning, and is best-in-class at identifying documents most likely to be miscoded by reviewers - meaning fewer iterative training cycles and cost to arrive at a higher precision level than competitive tools.

Client Case Results

- 1.7 million doc corpus
- 92% Recall, 70% Precision
- 72% of corpus eliminated as non-responsive
- **\$1.2 million estimated savings over linear human review**