iCONECT® V10 ANALYTICS AND CONTINUOUS ACTIVE LEARNING PRODUCT COMPARISON

An objective comparison between iCONECT Analytics and the 'TAR' offering of a leading eDiscovery tool.

WRITTEN BY JOSHUA GILLILAND, ESQ.

OVERVIEW

Attorneys conduct document review for many reasons in a lawsuit, whether it is conducting a reasonable inquiry to comply with Federal Rules of Civil Procedure Rule 11, preparing initial disclosures under Rule 26(a) to identify ESI supporting claims or defenses, or responding to a request for production under Rule 34. Whatever the reason for the review, there is a maxim in the law: Time is Money. The more time spent on document review, the more money is spent by the client for the lawyer's time.

iCONECT V10 Analytics, powered by Sentio Software™, looks to demonstrate a workflow that can save both time and costs.

THE CHALLENGERS:

In a recent product comparison between the following:

(a) iCONECT V10 Analytics, a leading eDiscovery platform who have recently embedded AI technology (with foundational code used in stock analysis) to create a seamless workflow for users.

(b) Product TAR, a major eDiscovery review platform with built in continuous active learning.

THE PROJECT:

Project X was a dataset from a previously reviewed case with 21,375 documents. The "documents" were electronically stored information that included emails, Excel files, and PowerPoint Presentations in native file format. This data was reviewed with both (a) iCONECT Analytics and (b) Product TAR giving the ability to directly compare results of both products from a common baseline document set. The review teams coded 75 records the same in each of (a) and (b) to train the respective advanced analytics.

THE TEST: (SUMMARY)

The product (a) vs (b) comparison was conducted by a client, who will be called Client Y, which is a Fortune 25 company involved in complex international projects.

The Project X teams began document review using the respective workflows for each of the databases to review and quality check first 100, and then another 200 documents, on four issue codes. The only difference was the technologies: the iCONECT Analytics and Product TAR.

UNDERSTANDING ICONECT ANALYTICS AND CONTINUOUS ACTIVE LEARNING (CAL)

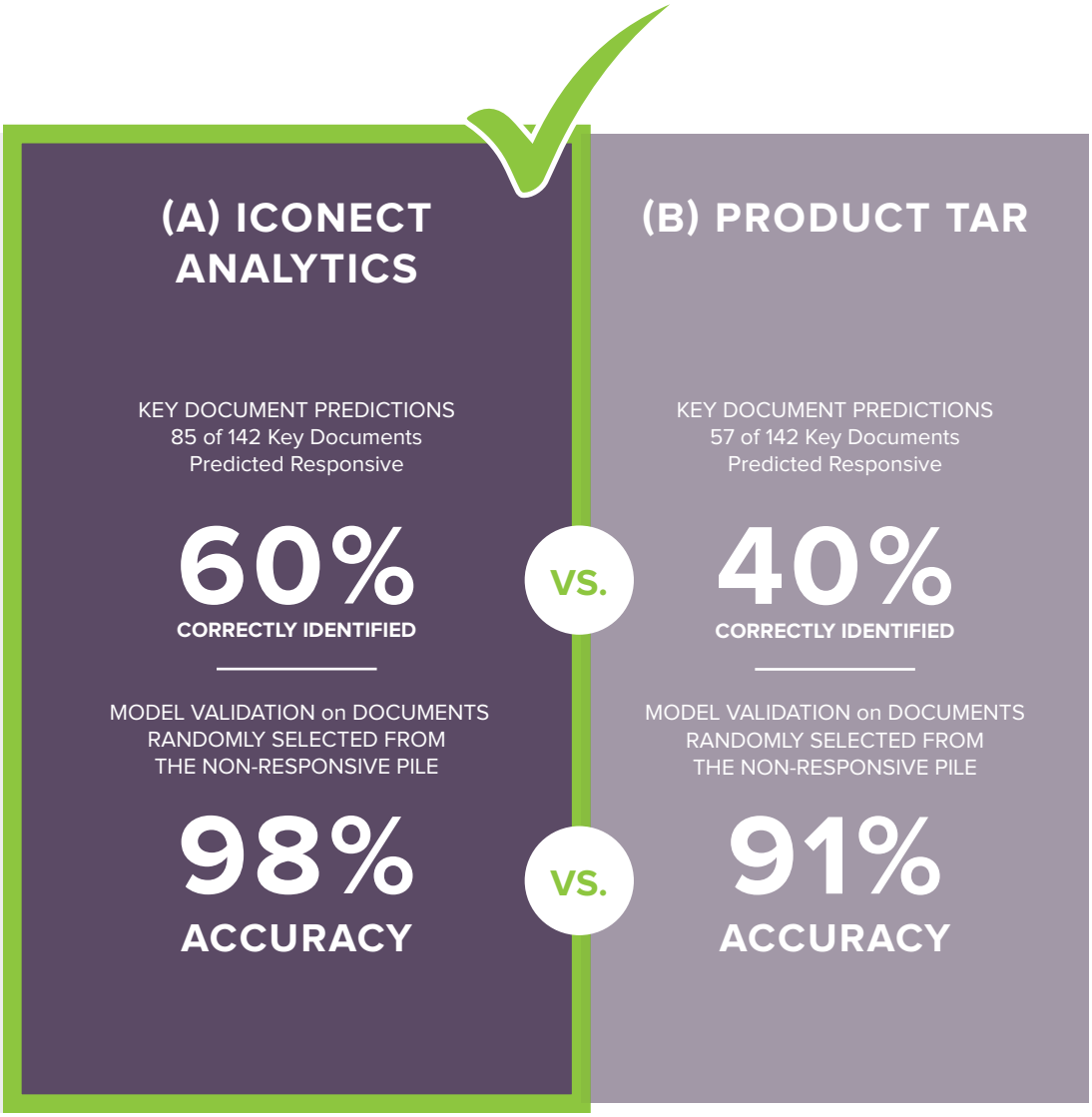
Many eDiscovery applications use a form of continuous active learning for predictive coding. Whether it is called technology-assisted review, or computer assisted review, this form of AI learns from the document reviewers' coding decisions. The Learning Index quantifies how well the predictive model is able to understand and predict the document universe. It is calculated by the number of documents it successfully predicts divided by the total number of documents.

In comparison, iCONECT Analytics creates a model by selecting a variety of documents to build the most accurate and comprehensive model. Each document reviewed is applied to the model building. Moreover, iCONECT Analytics uses a patented quality control (QC) process to ensure model building is based on accurate reviews. This helps to improve the accuracy of the model and yields better documents to review, which in turn reduces document review time.

THE RESULTS:

The results had iCONECT Analytics (a) correctly identify **60%** of the key documents predicted responsive compared to (b) predicting **40%** of the key documents. Moreover, the model validation test results on documents randomly selected from the non-responsive data had **98%** accuracy compared to 91% for (b).

The product comparison also found that iCONECT Analytics had a higher Learning Index than (b). In the first issue for coding, iCONECT Analytics had a Model Distribution Score of **.76** compared to .14; the second issue a score of**.84** to .02; the third issue **.88** compared to .03; and **.70** to .04 for the fourth issue.



iCONECT ANALYTICS FEATURES COULD HAVE INCREASED EFFICIENCY EVEN MORE:

CODE MULTIPLE ISSUES AT ONCE

iCONECT Analytics increases speed and efficiency in document review by applying up to twenty different issues against the document universe at once. Based upon the coding decisions made during the creation of the prediction models, the reviewers will be provided the best document to continue building each model. Furthermore, iCONECT Analytics is learning from every coding decision as they are made. For example, if the review team is coding for four issues each time issue code 1 is selected as responsive, the document has been coded as non-responsive for the other three issues. This allows multiple models to be built at once and using the iCONECT Analytics model refinement QC check the models are constantly updated and lead to a decrease in the time and amount of documents that need to be reviewed. iCONECT Analytics also uses proprietary heuristics to enhance the model building process.

VISUALIZE YOUR DATA IN CLUSTERS

Lawsuits have causes of action, which should be reflected in the issue coding. In a product defect case, there are claims for the defective production, and possible other causes of action, such as fraud in marketing, or breach of a warranty. These causes of action should be identified for issue coding, opposed to simply coding everything "relevant." This ignores how the data is relevant to the lawsuit and would create an extremely broad model. By focusing on the merits of the lawsuit, issue tags can be created for EACH cause of action. As this is applied in iCONECT Analytics, the respective models are applied to the document universe, the AI identifies clusters of documents with similar content for coding and near duplicates. These records could be email strings or different versions of a contract. While slightly different, they are likely relevant to the same issue in the lawsuit.

DON'T HAVE THE PERFECT DOCUMENT? CREATE ONE!

iCONECT Analytics can build a more efficient training set using the Xmplar functionality. While not done in the product comparison, virtual documents can be created to represent an ideal "smoking gun document," that can include priming terms or referencing pre-identified documents which represent each issue. Sorting the dataset by using email threading, near deduplication (such as contract that is both as a draft Word Document and executed PDF), or clustering similar records, can further enhance the training of the prediction model.

To put it simply, iCONECT Analytics provides reviewers documents similar to the ones they are reviewing based on their coding decisions. Instead of a static batch of documents for review, attorneys can focus their energies on what is relevant to the case. This is because iCONECT Analytics reindexing the database on the fly to train the AI on what is relevant based on coding decisions. This requires identifying the issues for review and conducting document review based on what supports the claims and defenses in a lawsuit.

DOCUMENT REVIEW COST SAVINGS BY THE NUMBERS (CASE STUDIES)

In a world with billable hours, efficiency is key to controlling costs. The results are not isolated to this one comparison. Many other projects have derived similar time and cost savings including the following case studies where the iCONECT Analytics engine, powered by Sentio Software, saved clients thousands of dollars from not reviewing irrelevant information:

CASE (I)

Corporate counsel used the iCONECT Analytics engine to prepare electronically stored information for trial counsel to use in the lawsuit. The lawyers reviewed 3,270 of the 677,653 records in the case (which was .5%). 634,350 documents were identified as non-responsive by the iCONECT Analytics engine. The system identified 43,150 documents with a confidence score of 80% as responsive for trial counsel to review. This saved 12,353 hours in review time and $573,746.

CASE (II)

In a medical malpractice case, the defendants produced 317,353 records to the plaintiffs. Leveraging the iCONECT Analytics engine to first issue code 1,305 records, the system identified 288,792 as non-responsive. Not reviewing the irrelevant data saved 5,900 hours and over $1 Million in document review.

CASE (III)

In a Foreign Corrupt Practices Act investigation with a dataset of 804,661 documents written in English, Spanish, and a combination of both languages, 808 records were issue coded for responsiveness. The iCONECT Analytics engine was able to identify 382,237 as irrelevant. This saved 14,332 hours in review time and an estimated $1.5 Million in attorney time.

Attorneys have a duty to secure the ***"just, speedy, and inexpensive determination of every action."*** [1] The iCONECT Analytics engine can help attorneys meet that duty by focusing attorney review on what is relevant to a lawsuit and not costing time and money to review irrelevant data.

THE POWER OF TECHNOLOGY: THE DUTY TO CONDUCT A REASONABLE SEARCH

Parties can seek discovery that is relevant to any party's claim or defense and proportional to the needs of the case. [2] Courts have recognized that using technology-assisted review such as predictive coding has "emerged as a far more accurate means of producing responsive ESI in discovery than manual human review of keyword searches." [3] Moreover, courts have stated that the "responding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information." [4] Furthermore, the main question for producing ESI is one of reasonableness and not perfection. [5] However, if the requesting party wants to challenge the sufficiency of a production, they must prove the production is unreasonable or inadequate. [6]

Consider a case where the defendant brought a short form discovery motion to compel the production of "the complete methodology and results of [Plaintiff's] TAR process" on the last day of fact discovery. The defendant made the "bald assertion" that the TAR information was necessary to "assess the adequacy of Entrata's [2] document production, as well as Entrata's document collection and review efforts." The Defendant did not provide any evidence the Plaintiff's production was deficient or question the adequacy of their document collection or review methodology and the court denied the motion. [7]

What this tells us is that courts recognize advanced analytics are more effective than humans reviewing ESI, that the litigants are in the best position to select the review strategy for their case, and that proving a production is inadequate requires far more than merely claiming a production is inadequate. In order to prove a production is inadequate, requesting parties have to show facts, which can be gaps in the production, such as date ranges, or specific individuals whose data is not included in the electronically stored information.

[1] *Federal Rule of Civil Procedure Rule 1.*
[2] *Entrata, Inc. v. Yardi Sys., 2018 U.S. Dist. LEXIS 104171, at *3-4 (D. Utah June 20, 2018), citing Rule 26(b)(1).*
[3] *Progressive Cas. Ins. Co. v. Delaney, No. 2:11-CV-00678-LRH, 2014 U.S. Dist. LEXIS 69166, at *8 (D. Nev. July 18, 2014).*
[4] *Hyles v. New York City, No. 10CIV3119ATAJP, 2016 U.S. Dist. LEXIS 100390, at *3 (S.D.N.Y. Aug. 1, 2016).*
[5] *Chen-Oster v. Goldman, Sachs & Co., 285 F.R.D. 294, 306 (S.D.N.Y. 2012), citing The Sedona Conference, The Sedona Conference Database Principles: Addressing the Preservation and Production of Databases and Database Information in Civil Litigation, March 2011 Public Comment Version, at 32.*
[6] *See generally, Terry v. Register Tapes Unlimited, 2018 U.S. Dist. LEXIS 50846, at *6-7 (E.D. Cal. Mar. 26, 2018).*
[7] *Entrata, Inc., at *10.*

DOCUMENT REVIEW WITH iCONECT ANALYTICS

Litigants must conduct a reasonable search to find responsive electronically stored information and conduct their cases cost-efficiently. Lawyers can meet these duties using iCONECT Analytics to quickly find responsive data and avoid unnecessary costs reviewing irrelevant ESI.

ABOUT iCONECT

iCONECT Development, LLC develops the innovative iCONECT eDiscovery review software platform. iCONECT raises the bar by delivering intelligent, easy-to-use tools that help hosting providers, law firms, and legal departments optimize workflows and manage some of the world's most complex legal cases more efficiently. Leading AI and auto-redaction capabilities combined with a user's ability to search, sort, analyze, categorize and produce documents and multi-media files recently led industry publication 'Silicon Review' to name iCONECT as one of the '30 Fastest Growing Tech Companies' of the year.

iCONECT Analytics is powered by Sentio Software www.sentiosoft.com

Contact iCONECT for more information.
www.iconect.com | info@iconect.com | 1-855-915-8888

ABOUT THE AUTHOR

Joshua Gilliland
Twitter: @bowtielaw
Josh@bowtielaw.com

Joshua Gilliland is a California attorney creator of the eDiscovery blog Bow Tie Law and has presented at over 400 eDiscovery seminars and webinars. Josh is co-creator of The Legal Geeks blog and podcast, which has made the ABA Journal as one of the top 100 blogs for lawyers from 2013 to 2018. Josh grew up in Silicon Valley and is a graduate of UC Davis with a degree in Political Science and earned his law degree from McGeorge School of Law, University of the Pacific. Josh enjoys organizing panels and mock trials at comic conventions, photography, and volunteering in Scouting.