



An interview with Craig Carpenter of Recommend: a discussion on predictive coding

Feb 2nd, 2010 | By Gregory P. Bufithis, Esq.

This interview is part of our new series “Data! Data! Data!” — Cures for a General Counsel’s ESI Nightmares”. For our introduction to the series [click here](#).



Craig Carpenter, Vice President of Marketing for Recommend, oversees all aspects of marketing at Recommend. He has extensive experience in the enterprise software, information security, and eDiscovery industries, and is a frequent speaker and panelist at industry events. He is also an adjunct faculty member at the University of San Francisco, where he has taught graduate classes on high tech marketing, content management and digital rights management (DRM).

We caught up with Craig at this year’s LegalTech.

TPL: Recommend won top honors in the 2009 Law Technology News Vendor Awards. The readers of Law Technology News selected Recommend’s MindServer Search as the best product in the Knowledge Management Systems category. Our congratulations.

CC: Thank you. We have always worked closely with the legal industry, so that award means a lot to us. Legal professionals work in one of the most challenging environments for enterprise search. They need to locate precise information under ever-shorter deadlines, with ever-increasing stakes. Many need to locate documents across multiple offices, often in different countries with different languages. So we developed our solutions in close collaboration with law firms and legal professionals to address their needs. It was wonderful to have LTN readers validate those efforts.

TPL: Give us an overview of your MindServer Search product.

CC: MindServer Search is our flagship enterprise search product. It provides highly accurate and relevant search results through a simple, intuitive interface. It uses proprietary, machine-learning technology to automatically create concept models based on the information within the enterprise. That gives it the unique ability to accurately identify and rank relevant information for each user without the need for additional input from the user.

For our legal customers, we also offer a popular Matters and Expertise module for MindServer Search, which enables them to find all relevant matter information and expertise within the firm. The module's Expertise Location feature automatically updates areas of expertise based on work product, projects, clients, etc., which makes it simple to find attorneys with relevant experience on a particular topic, as well as the documents and matters associated with them.

TPL: And it's built on Recommind's CORE™ (Context Optimized Relevance Engine) platform?

CC: Correct.

TPL: Which is...

CC: The CORE platform is the conceptual search and automatic categorization engine that powers all Recommind products. It identifies distinct concepts within a document or data set, regardless of the language or subject, and organizes and prioritizes all documents accordingly. The system is fully automated, so it does not require any thesauri, dictionaries or ontologies – although it can certainly work with them. CORE is what makes it possible for our search and eDiscovery software to deliver such extraordinary accuracy and automation.

TPL: At the IQPC eDiscovery event last month you said that Predictive Coding™ technology was now accurate enough to take the place of contract reviewers. Explain.

CC: There's no doubt Predictive Coding is accurate enough – this has been proven in many cases. A number of AmLaw 30 firms have proven it by using Predictive Coding and comparing it to the results from contract attorney review (and partner review as well) on the same data. The results in every case were that they achieved better accuracy with Predictive Coding, and in the process saved 50-80% of what they would have spent on traditional review because contract attorneys were either not needed or were able to work far more efficiently (or both). This is what we mean when we talk about revolutionizing the economics of eDiscovery; no one else is doing this.

TPL: But is the technology fully defensible? I am thinking of random sampling of the automated output data I have seen.

CC: Absolutely. It's as defensible, or more so, than traditional review, thanks to random sampling of the output and the robust methodology that is part of Predictive Coding. Technology is only part of the solution; the rest comes in the form of processes, workflows and methodologies which can only be devised in conjunction with customers (hence our close

collaboration with our customers). In traditional review, for a number of reasons you don't necessarily go back and sample the decisions made by the attorneys, which carries with it a certain amount of risk. With Predictive Coding you do, and you're able to measure the accuracy of the review, both human and computer. It's a not-so-well-kept secret that linear review is typically not an overly accurate or consistent way to go about things, but it has been the devil we know. Predictive Coding is a quantum leap in document review efficacy and accuracy.

TPL: And judges have demonstrated that they consider automated review acceptable?

CC: In effect they have. Judges care about getting to a just result as efficiently as possible; they care far less about the means used to get to that result – so long as the means do not undermine the pursuit of justice. So judges are not in the business of “validating” any particular technology or process. That said, given the broken economics of today's eDiscovery judges have definitely been expressing a fervent desire for a better approach to document review, and prominent judges like Judge Facciola, Grimm and Peck have indicated that technology can and should be brought to bear on the problem, because it can really help. It's important to look at how the top litigation firms have responded now that they have a mandate from judges to change the economics of eDiscovery. And if you look at the top firms in the world — WilmerHale, Morgan Lewis and Fulbright & Jaworski, just to name three — they have made a commitment to Predictive Coding as the future. That's a very, very strong endorsement.

TPL: So whereas early case assessment (ECA) was the mantra in 2009 —winnowing relevant data down to reduce the number of documents to review — you think Predictive Coding is the “new new thing” for 2010?

CC: Yes, we think this is the year Predictive Coding will start replacing linear review left, right and center. Everything is in place. The technology is capable, defensible and in use by leading firms. And given that it cuts document review costs by 50%-80%, it's hard to justify not using it.

TPL: So we are headed down the path to where machines can be statistically proven to be as accurate as human review? Is the technology getting to the point where we can also winnow out the eyeballs — contract attorney reviewers? No room for a human element and perspective?

CC: There will always be a need for human review in eDiscovery. But bear in mind that the traditional eDiscovery process relies on an outdated, paper-based model that requires attorneys to sit in a room and review terabytes of ESI, one at a time. That's a textbook example of work that should be assisted by intelligent automation. With the continuing rise of eDiscovery, there will always be plenty of work for attorneys. Some firms, and some clients, will always want to have an attorney's eye on every document – which does not at all preclude the use of Predictive Coding. Even in such a case, they can perform that task much faster and more consistently using Predictive Coding.

TPL: So you see automated review as a revolutionary, not evolutionary change for eDiscovery.

CC: When you consider how much most lawyers dislike change, it's revolutionary, no question. Up until now eDiscovery tools have aimed to cull data down, which is a pretty small step in the evolutionary direction. Predictive Coding, by contrast, is a huge leap forward because it completely overturns the way document review has been conducted – which is by far the most expensive part and thus the root cause of the disease. Most estimates put review costs at 60-70% of the total cost of eDiscovery. Every other approach has tinkered around the edge of the problem. With Predictive Coding, we have fundamentally and permanently changed the economics of eDiscovery.

TPL: Last year we had an opportunity to see a demonstration of your product Insite Legal Hold, which is an ECA product providing processing, culling, collection and hold. It seemed to literally “crawl” and index information from the many data sources a company has. I assume it allows a company to conduct “live” and in-place exploration of data before applying a legal hold to quickly and efficiently locate and preserve potentially relevant data. Tell us about it.

CC: Sure. Insite Legal Hold is our all-in-one software for Early Case Assessment, identification, preservation, collection and processing. And you're right, one of its most distinctive features is its

Explore-in-Place technology, which enables users to find, sample and assess ESI before collection. That allows enterprises, number one, to conduct ECA as early as possible — which is the whole point with ECA — and number two, only collect the ESI they need to collect for any given matter. This far more targeted and precise approach in turn significantly lowers eDiscovery costs and makes them identifiable and predictable from the very outset of a matter — which is a “game changer” for our customers. It works hand-in-glove with our Axcelerate eDiscovery software for Predictive Coding and deep analysis. Together, the two provide the most accurate and cost-effective eDiscovery solution available.

TPL: We noted that Law Technology News also awarded Reed Smith LLP a gold medal for its intranet portal ouRSpace which utilizes Recommind's MindServer Search technology. How did that come about?

CC: We've known Tom Baldwin for years. When he left Sheppard Mullin to come to Reed Smith, he had a vision of what he wanted to do to create a next-generation information system and infrastructure, which was highly automated and highly intelligent, and was able to deliver the appropriate information to users when they needed it. We sat down with him to talk about his vision. It was a highly collaborative process where we helped him learn from what all of our other law firm clients have done—a benefit of being the market leader—but Tom had a terrific vision for where he wanted it to go. And now Reed Smith is reaping the rewards of his foresight and the collaborative effort. Their lawyers and staff have a significant competitive advantage because of their ability to access the exact information and expertise they need when they need it.

TPL: You have some other law firms on board [Morgan Lewis, Sherman & Sterling, WilmerHale, etc.] and we interviewed Tess Blair at Morgan Lewis as part of this series. You also have some stellar enterprise clients: Marathon Oil, Bertelsmann, BMW, Verizon, etc. Is your client base equally split between in-house and law firms, or is it weighted toward in-house?

CC: It's more evenly split. The two are in many cases interrelated with each other, certainly on the eDiscovery side. What we're finding is that big legal departments and law firms have a similar need for secure access to relevant information with search, and in eDiscovery, they both need to change the economics of eDiscovery and hone in on key documents while quickly separating out the noise. That's what we offer to law firms and their corporate clients. And because we can deliver on that promise, we've had a great deal of success with both sets of clients.

TPL: We had the opportunity last year to speak with Hartwig Laute, Director European Operations for Recommind. He showed us the Bundeskartellamt use case (the German competition regulator, akin to the US FTC) and explained how and why the German Cartel Authorities rely on Recommind technology. Tell us about your European operations.

CC: Recommind has had European operations since its inception, and they're very much an integral component of what we do. Roughly half of our employees are in Europe because that's where we do much of our development work. Our CTO/co-founder Jan Puzicha is in Germany, and we also have significant operations in the UK. The majority of our clients are in the US, especially on the law-firm side, but our business is growing on both sides of "the pond." Most of the large law firms are here or in the UK, so we work with them in those countries, but on the enterprise side we have a number of clients across Europe as well as in the United States.

TPL: E-discovery costs are skyrocketing. Yet much of EDD is now a commodity – and that has changed the structure of the market. Prices are — shall we say — more predictable and probably more realistic.

E-discovery vendors have capped fees, set flat fees or worked with various forms of pricing estimators. Have you changed your pricing?

CC: We have certainly added more choices to our price list to accommodate the overwhelming demand we've seen, but if you are asking if we have had to lower our prices the answer is not at all. It's definitely the case that much of the eDiscovery process, including culling, processing, hosting and forensic imaging, has been commoditized; older vendors trying to maintain market share and the rather simplistic appliance offerings and vendors have pushed this trend. But where we play and what our products are capable of doing for clients – Predictive Coding being perhaps the best example – is nowhere near becoming commoditized. The basic problem with eDiscovery is that it still uses the paper-based, linear review model, even though 99% of information these days is digital. Most EDD products try to alleviate the symptoms of that problem rather than address the problem itself, e.g. "better" linear review, a simple culling appliance, etc. Those technologies are commoditized now or will soon be. But we attack the fundamental problems of eDiscovery, the illness rather than its symptoms. Predictive Coding doesn't just streamline document review for human reviewers—though it delivers that too—it actually automates the majority of the process using intelligent technology and defensible workflow. That's something no other company or technology can deliver – period. And because it truly is game-changing technology, law firms and clients alike are more than willing to pay a premium. After all, it will save them a tremendous amount of time and money so the investment is easy to justify. Because this is so unique and such a difficult problem, in spite of a noisy market there's no danger those capabilities will be commoditized any time soon.

TPL: Craig, we greatly appreciate your time.

CC: Thanks very much. We're regular Posse List readers, and it's always a pleasure to talk with you.

Postscript: You can find Recommind in the Exhibit Hall at Booth #2104 and 2106.

Gregory P. Bufithis is the founder and chairman of The Posse List and its sister sites The Electronic Discovery Reading Room (<http://www.ediscoveryreadingroom.com>) and The Posse Ranch (www.theposseranch.com). He is also founder and chairman of Project Counsel (www.projectcounsel.com).