

Document Imaging Report

Business Trends on Converting Paper Processes to Electronic Format

4003 Wood Street ● Erie, PA 16509 ● PH (814) 866-2247 ● <http://www.documentimagingreport.com>

December 4, 2009

THIS JUST IN!

DIR INTRODUCES COMPLIMENTARY NEWSLETTER

You may have already seen it, but *DIR* has launched a complimentary e-mail newsletter designed to reach a broader community than the vendors, systems integrators, and high-end end users we typically serve. The e-mail newsletter does not feature the same premium content as our regular newsletter, but will feature an occasional article pulled from the pages of *DIR*.

It also features links to edited versions of imaging-related news stories from the previous weeks, and we've left room for guest contributors as well. In addition to our regular subscribers, this complimentary newsletter is being distributed to contacts I've met through events, as well as people who have attended Webinars I've given and signed up for free trials to *DIR*. A lot of these people don't fit the profile of a reader of our regular newsletter, but are obviously interested in the market. Our plan is to offer the newsletter as a marketing vehicle to industry vendors, so please contact me if you're interested.

Sean Parnell, president of **Innovaxis, Inc.**, is helping with the development of the newsletter. Sean is a former Böwe Bell + Howell executive who has also done work with Techinfocenter. Innovaxis is Chicago-based consulting firm offering services in areas like development of strategic marketing plans, the launch of new products and services, designing Web sites, and developing marketing collateral.

For more information:
<http://www.innovaxisinc.com/>

Vendors Position Themselves For Mid-Market IDR Growth

ReadSoft signs deal with Altec; Datacap releases FastDoc.

Invoice capture has half-jokingly been referred to as a horizontal/vertical application. This means, while it's focused on a specialized area—accounts payable (A/P), it still can be applied in any business—as everyone has A/P processes. That's one reason invoices were chosen as the initial application for intelligent document recognition (IDR) technology—everybody has them.

Over the past few years, we've seen fairly good progress of automated invoice capture technology in the Global 2000 market. ISVs like **ReadSoft** and **Brainware** have won some well-publicized six- and seven-figure deals. Other vendors like **Kofax**, **AnyDoc**, **Datacap**, **Paradatec**, **Top Image Systems**, **ABBYY**, **BancTec**, and **Captiva**, have also announced impressive high-end wins. And while the price tags on these large deals are often impressive, the fact that the market is defined by a number like "2000" means it is limited. [That's not to say it's close to being saturated. At Harvey Spencer Associates' Document Capture 2009 conference, **Paystream Advisor's** Henry Ijams described IDR technology as being in the "innovator/early adopter" stage in A/P departments. He also reported that only 16% of companies with more than \$500 million in revenue use OCR on more than 40% of their invoices, see *DIR* 9/18/09].

In contrast, the SMB (small-to-medium-sized business) market offers an almost unlimited number of opportunities. And this is where many invoice capture pioneers have begun to look for additional business. I ran a feature discussing some of these downstream initiatives after I saw several of them demoed at this spring's AIIM show [see *DIR* 4/3/09]. And recently I caught up with executives from ReadSoft and Datacap—both of which have announced new initiatives targeting the SMB. [I'll offer one more qualifier before moving forward. This is certainly not the first article I've done on mid-market solutions for automating invoice capture.

But, much like I saw in the high-end of the market a few years back, it seems the early hype is outpacing adoption rates. However, as adoption rates did eventually catch up in the upper end, I'm expecting it's just a matter of time before SMBs start adopting IDR at a similar rate.]

ReadSoft to leverage Altec channel

One challenge to marketing to the SMB segment has always been defining it. Fortunately, in the A/P space, SMBs can most easily be identified by the type of accounting software they're using. ReadSoft, which has had success in the high-end through its integration with **SAP** and **Oracle** systems, has chosen to go to market in the SMB space through a partner that focuses on integration with SMB accounting apps.

ReadSoft, which is based in Sweden and has its U.S. headquarters in the New Orleans area, recently announced a worldwide distribution agreement with **Altec**. Altec is a Laguna Hills, CA-based ISV that has been featured in *DIR* several times. Originally, a vendor of printed forms specializing in servicing **Epicor** accounts, Altec acquired a document imaging software developer in 2001 and has built a successful business image enabling mid-market ERP and accounting systems. To date, Altec has not offered any OCR. Instead, it has relied on its unique *Smart Form* technology to reduce key entry of data [see *DIR* 3/2/09].



Brandt Morrell,
president and COO,
Altec

"Last year, we started hearing a lot of noise from our customer base asking for OCR and IDR technology that we could plug-in to the front end of our *doc-link* application," said Brandt Morrell, president and COO of Altec. "We spent about eight months searching for a leading OCR player whose product we could bring to our market." [Altec also may be feeling pressure from U.K.-based **Version One**, which recently opened a U.S. office. Version One targets **Sage** resellers (also an Altec market) and resellers of other mid-market ERP/accounting systems. Version One has been enjoying success in the U.K. with an OEM version of **ABBYY's** IDR technology.]

"To date, nobody has really found a way to sell automated invoice capture successfully in the mid-market," said Mark Tague, CFO and CTO of Altec. "Over the years, our customer base has questioned us about it, but we have skirted the issue. We were not very well educated on OCR and IDR, and we had not embraced it. We worked with third-party integrators and had some success. But, to be really successful, we felt our customers needed one throat to choke.

"The partnership with ReadSoft will enable us to do things like set correct expectations and offer the same level of support and intimacy for capture as we do for our *doc-link* implementations. We will be selling the ReadSoft software, doing the implementations, and providing part of the support.

Document Imaging Report

Business Trends On Converting Paper Processes To Electronic Format

DIR is the leading executive report on managing documents for e-business.

Areas we cover include:

1. Document Capture
2. Image Processing
3. Forms Processing/OCR/ICR
4. Enterprise Content Management
5. Records Management
6. Document Output
7. Storage

DIR brings you the inside story behind the deals and decisions that affect your business.

Vol. 19, No. 23



Editor: Ralph Gammon
4003 Wood Street
Erie, PA 16509
PH (814) 866-2247
FX (412) 291-1352
ralphg@documentimagingreport.com

Managing Editor:

Rick Morgan
PH (814) 866-1146
rickm@scandcr.com

DIR is published 24x per year, on the 1st & 3rd Fridays of the month, by:

RMG Enterprises, Inc.
4003 Wood Street
Erie, PA 16509
PH (814) 218-6017
<http://www.documentimagingreport.com>

Copyright © 2009 by RMG Enterprises, Inc. Federal copyright law prohibits unauthorized reproduction by any means including photocopying or facsimile distribution of this copyrighted newsletter. Such copyright infringement is subject to fines of up to \$25,000. Because subscriptions are our main source of income, newsletter publishers take copyright violations seriously. Some publishers have prosecuted and won enormous settlements for infringement. To encourage you to adhere to this law, we make multiple-copy subscriptions available at a substantially reduced price.

Subscriptions: \$597 (electronic) or \$670 (paper) per year.

"We also have tremendous reach and relationships with accounting ISVs and VARs. They know Altec and trust us. These relationships helped us successfully bring to market our document imaging product after we acquired it. We hope to enjoy similar success with the ReadSoft product."

Altec has more than 1,200 customers under maintenance, as well as an established channel of 250 North American resellers. Altec also has a growing international presence. "In addition to its channel, we are impressed with the technical integrations Altec has with mid-market products from Epicor, Sage, and **Microsoft** [the *Dynamics* line]," said Bob Fresneda, president of ReadSoft North America. "We've done well in the high-end of the market by offering our own workflow technology. In the SMB market, our strategy is to only provide the capture component and let ISVs like Altec provide the workflow."



Mark Tague, CFO and CTO, Altec

Final details of the agreement were still being worked out when we talked, but Fresneda is confident that ReadSoft will have no problem scaling down its offering. "We are not publicizing our pricing, because we don't want the rest of the market to know it," he told *DIR*. "Working with Altec, we've learned a lot about how to price for their market. Our pricing actually goes down to accommodate as little as 5,000 invoices per year. In Europe, we are selling implementations for as low as 2,000 invoices per year."

"We've figured out how to package a solution that best targets the mid-market. Our product has different modules, some of which are better to start with for Altec's customers and channel. For example, we're not initially rolling out line-item capture to the mid-market."

"Right now, we think that more than half of Altec's customers are potential ReadSoft customers. However, it's our goal to continue to evolve so that all of Altec's customers can use our software, not only for invoices, but for other types of documents as well."

Morrell said Altec will initially be targeting A/P departments with at least 20,000 invoices per year. "Once you reach the point where you have three to four heads down A/P processors, automated data capture starts to become cost justifiable," he said. "We estimate that with *doc-link*, our customers can

increase their productivity 25-40%. When you lay down ReadSoft's OCR capabilities on top, it can push that improvement to greater than 50%."

Tague added that introducing automated data entry should help increase overall sales of *doc-link*. "Now, we can really offer an end-to-end A/P processing solution," he said. "In the past, we'd get objections that, even though potential customers could scan documents, they'd still have to manually index them. A lot of those objections will go away."

Brandt said the ReadSoft technology will work in conjunction with *Smart Forms*. "*Smart Forms* basically enable our customers to do their invoice coding, data entry, and approval outside their ERP systems," he said. "For invoices that have POs, a lot of the information they need can be downloaded from the ERP system, but the user still has to enter a PO number to trigger the population of the *SmartForm*. ReadSoft's technology will help automate that step."



Bob Fresneda, president, ReadSoft North America

Fresneda indicated that ReadSoft is looking forward to being able to leverage its reputation to assist Altec's sales efforts. "Because we're perceived as the market leader, we are contacted quite often by mid-market companies looking for an invoice capture solution," said Fresneda. "In the past, we didn't offer a solution that could address that segment. Now, we can pass those leads on to Altec."

FastDoc has wide-ranging potential

Like ReadSoft, Datacap is looking to develop a new channel to sell its mid-market solution. However, *FastDoc Capture*, which was previewed at AIIM 2009 and released for general availability this week, is aimed more at the "S" in the SMB than the ReadSoft-Altec solution. It carries a list price \$2,995 for a single client (there is no server component).

"*FastDoc Capture* is designed for the companies that attend ITEX [an event for copier dealers] and want to get into imaging," said David Jenness, marketing manager for Datacap. "It's also OEM-friendly and should appeal to a broader base of resellers than *Taskmaster* [the product for which Datacap is best known]. *Taskmaster* is more robust, but it also requires a lot of training and technical savvy. [Also, an average deal size for *Taskmaster* is also around \$80,000]."

Jenness stressed that *FastDoc* is not restricted to invoice capture. "Yes, it can do invoices, but if you

want a full-featured A/P automation application, you're better off with *TaskMaster*," he said. "*Fast Doc* can be used for applications like human resources and medical documents, as well as backfile conversions. We had one of our beta customers using it to capture certificates of insurance in a property management application.

"We have some ideas of how we think *FastDoc* will be deployed. However, sales over the next six months will really tell us where the action is. Right now, our current channel and OEM partners are making the most noise, and each one has a different focus. In a few months, we should have a better story about where *FastDoc* will play in the market."

In fitting with its target market, *FastDoc* is designed to be easy to install, set-up, and use. When setting up *FastDoc*, users define document types and indexing fields they want to capture for each type. A user also sets up the export destinations for each document type—either a system folder or a *SharePoint* library.

FastDoc features a TWAIN scanning interface and can also import files from a folder—which could be a destination for MFP scanning. *FastDoc* performs image clean-up and gives users the ability to do manual separation or utilize bar-coded separator sheets. It then performs full-text OCR and offers some advanced data capture capabilities.

Basically, the first time *FastDoc* sees a particular type of document, say an invoice from ABC Vendor, the user tells the application it's an invoice. This prompts *FastDoc* to display the pre-set meta data fields, such as "date," "invoice number," "P.O. number," and "amount." The user then clicks on a field and uses the mouse to draw a box around the area on the image which shows the data that corresponds to that field. [You can watch a demo of all this at <http://www.datacap.com/products/fastdoc/>.]

Rules can be set up for validating the format of the index data and even for comparing it to information that has been downloaded from databases and stored in tables. Key entry can be used to correct any data misinterpreted by OCR.

The next time *FastDoc* encounters a similarly set-up invoice (potentially from the same vendor), it will automatically classify the document and capture the meta data. The classification process uses a combination of Datacap's digital fingerprinting

technology [which was originally developed for *TaskMaster*] and text captured through OCR.

After the meta data is captured, it is exported, along with a TIFF or PDF version of the image. The meta data can be released in a variety of formats to accommodate ingestion into third-party apps.

"While we've traditionally focused on the high end of the market, we couldn't help but notice that the mid-market presents some interesting opportunities,"

said Jenness. "We think there is a sweet spot between what Harvey Spencer defines as the batch capture and ad hoc segments. If we can deliver the efficiencies of traditional batch capture to ad hoc users,

we think we'll have something pretty special. We are offering mid-market users—who are increasingly doing scanning with MFPs—automated separation, classification, and the opportunity to reduce manual data entry."

According to Tom Simalchik, Datacap's product marketing manager, setup for *FastDoc Capture* can be completed in five minutes. "We have done some benchmarking tests, which have given us a sense of how many documents per day *FastDoc* will be good for," he added. "The license is for an unlimited number of pages and fields, and we've tested it on several thousand documents per day.

"Most of the training for the product is Web- and video-based. There are some integration opportunities for resellers who want to customize data and image output so it can be ingested more readily by an ECM system. That said, we are very excited about our *SharePoint* integration. As the market for document management solutions built on *SharePoint* matures, we think we'll have a pretty attractive offering for *SharePoint* integrators."

Bob Lastorka, Datacap's director of business and channel development, concluded by saying *FastDoc* capture should open up a host of new opportunities for Datacap. "Our OEM partners are clamoring for it immediately," said Lastorka. "We are adding it to every contract. A lot of our partners' users are intimidated by *Taskmaster*. They might focus on image-enabling *Sage* or *Microsoft Dynamics* accounts and their customers aren't used to supporting enterprise-class applications.

"We've also received interest from vendors that sell applications into IBM's mid-size iSeries

"If we can deliver the efficiencies of traditional batch capture to ad hoc users, we think we'll have something pretty special."

—David Jenness, Datacap

environments. We are in discussions about replacing some existing capture offerings. Some partners are even talking about putting their own products together that combine both *Taskmaster* and *FastDoc*."

For more information:

<http://www.altec-inc.com/>; <http://www.readsoft.com/>
<http://www.documentimagingreport.com/Altec-ReadSoft.1748.0.html>
<http://www.datacap.com/news/headlines/archives/fastdocgr/>;
<http://www.datacap.com/products/fastdoc/>

Capture ISVs Report Strong Quarters

The image capture software market appears to be bouncing back from the world economic downturn that affected everyone in late 2008 and early 2009. Over the past month, we received positive reports from a trio of capture software vendors: **Datacap**, **Top Image Systems** (TIS), and **Kofax**.

"The third quarter was amazing," said David Jenness, marketing manager for Datacap. "And the fourth quarter is shaping up pretty good as well. It looks like we're going to end 2009 with 5-10% revenue growth over 2008. That seemed very unlikely in June."

Jenness said Datacap continues to have success in the invoice capture market. "We continue to improve the product by listening to our customers," he said. "One feature we've introduced recently is purchase order line reconciliation. This makes it easier to do three-way matching [between invoices, POs, and shipping documents]."

Jenness said the company's OEM business is also picking up. "This includes a deal we signed recently with a major integrator in the healthcare space," he said. "We also continue to pursue logistics applications. We now have implementations with three leading companies that account for 22% of the logistics market."

Public companies report growth

Kofax and TIS each reported growth in the third quarter as well. Kofax CEO Reynolds Bish was quoted as saying in regards to the three months ended Sept. 30, "I'm pleased to report that we made good progress and performed better than expected in our software business this past quarter....We continue to expect low- to mid- single digit organic revenue growth in our software business this financial year."

As for TIS, it reported revenue of \$5.9 million for

the third quarter, which represented 4% growth over the previous quarter. "Our quarterly sales actually decreased 29% on a year-over-year basis, but we've eliminated a lot of hardware and other low-margin sales," said CEO Ido Schechter during a conference call for investors. "[Comparing apples to apples] our sales would have been up 5%. We also reported a gross margin of 61%, which was our highest in two years. This helped us generate a positive cash flow of \$250,000.

"We have emerged from the economic downturn as a leaner, more streamlined, and focused company. As market conditions improve, we expect to be ready to benefit and thrive. Geographically, we performed very strong in Europe and the U.S., which together were responsible for more than half our revenue in the third quarter."

Schechter concluded that TIS' sales pipelines are improving. "In some places, like the U.K., the pipeline is great, but we're not closing as many deals as we'd like. Our closing rate for our pipeline is currently about 20%, while in better economic times we expect it to be about one in three."

For more information: <http://tinyurl.com/TISQ3>;
<http://tinyurl.com/KofaxQ1>

Distribution Agreements Target Emerging Markets

As part of an effort to expand its global presence, **BancTec** recently signed four new distribution agreements. Three involve Middle Eastern partners and the fourth is with a Russian company. The deals are the work of BancTec Technologies, a business unit launched in the beginning of the year to expand the scanner and software vendor's presence outside of Europe and North America.

"In North America and Europe, we have been selling for a long time through a combination of direct and partner channels," said Mark Fairchild, BancTec's CTO and the president of BancTec Technologies. "In the rest of the world, we sell primarily through partners. We did some market analysis and determined the higher growth economies where we wanted to establish stronger relationships and distribution. These deals are the initial results of our efforts."

The Middle East distribution partnerships will cover countries like Saudi Arabia, Kuwait, Bahrain, Qatar, Jordan, and the United Arab Emirates. "Regardless of the recent news out of Dubai, the Middle East has some strong and growing economies," said Fairchild.

"We are seeing a number of opportunities related to e-government initiatives. Overall, there is an increasing desire in these countries to make more information available over the Internet."

Fairchild described Russian partner **PIRIT**, which is based in Moscow, as primarily a hardware distributor. Its Web site lists Fujitsu, Canon, Kodak, and Toshiba as partners. In addition to Russia, PIRIT sells in other former Soviet states like the Ukraine, Belarus, Armenia, and Georgia. "When considering the BRIC [an acronym for Brazil, Russia, India, and China—which are considered to have large, rapidly growing economies] countries, we already have relationships in the B, I, and C. Now we've established distribution in the R," said Fairchild. "Eastern Europe is another area we are looking at."

All the distribution agreements are initially focused on BancTec's hardware—its IntelliScan and E-Series high-speed document scanning and sorting transports. "There is the potential to introduce software in a subsequent phase with some of these distributors," said Fairchild. "It would have to be in an appropriate region and through a partner with the right integration skills".

BPO revenue rising

The new distribution agreements should help shore up BancTec's declining revenue from its hardware and software products. According to the company's most recent S-1, after peaking at over \$97 million in 2007, revenues from this portion of BancTec's business declined 24% in 2008 to \$74 million. Through the first six months of 2009, revenue was down another 10%.

In past discussions with BancTec executives, they've noted that some of this decline was expected as increasing adoption of Check 21 and electronic payments has reduced requirements for check sorting hardware. Fairchild noted that the new geographical markets open up new opportunities for BancTec's sorting equipment. "There are certainly checks in these countries, but there are also other types of check-sized documents that we don't have in the U.S. One example is a deal we just signed in Europe for a system to process food tickets. These are basically vouchers that companies provide to their employees to spend at the local restaurants."

Helping offset BancTec's declining hardware revenue has been an impressive increase in business process outsourcing (BPO) revenue. In 2007, BPO accounted for \$37 million of BancTec's revenue. BancTec surpassed this figure in just the first six months of 2009, during which it generated \$42.1 million in BPO revenue. This represents almost 50% growth over the first half of 2008.

"Our BPO growth is coming from multiple markets," said Chuck Corbin, VP, worldwide marketing, BancTec. "Clearly, with the condition of the global economy, companies are looking to be more efficient. BPO fits right in with this strategy. Because of our legacy, payment processing remains a strong segment for us. But, overall, we're processing a broad mix of documents. This includes invoices, tax forms, and more general types of forms."

According to Fairchild, BancTec's BPO revenue is fairly evenly split between the North American and European markets. "We may look at BPO opportunities in other geographical markets, but it's not something I'm ready to talk about yet," he said.

For more information: <http://www.pirit.com/>;
www.emitac.ae; <http://www.banctec.com/>;
<http://www.banctec.com/services/bpo-services/>

Brainware Announces New Features For Distiller

Document capture and IDR specialist **Brainware** recently announced several upgrades to its flagship product line with the release of *Distiller V5*. V5 features improved scalability, a zero-footprint verification client, and support for recognition of double-byte character sets. Brainware also recently introduced a module that enables vendors to self-check on the status of their payments.

According to Charles Kaplan, VP of marketing and product management for Brainware, the new version of *Distiller* can be run off SQL and **Oracle** databases. "Our legacy versions utilize a file system architecture," he said. "The new database architecture will improve *Distiller's* scalability, as well as its ability to run back-ups. It also offers better support for our Web client, as well as future product improvements we have coming down the pike. This will allow for some very robust capabilities going forward."

The zero-footprint client enables browser-based verification and correction of fields not recognized with acceptable confidence levels by *Distiller's* OCR engine. "The new Web client was originally developed to meet the requirements of a customer rolling out *Distiller* to multiple shared services centers around the globe," said Kaplan. "They wanted the feature set of our thick client, but also wanted a clean IT roll-out. They didn't want to manage ActiveX downloads that might disrupt their standard operating environments."

"We leveraged .NET and AJAX technologies to create a zero-footprint verification client. It's very robust. It includes features like overlays, the ability to

suggest field-level substitution candidates, drag-and-drop rubber-banding, and the ability to pre-populate fields based on table data.”

Kaplan said the zero-footprint verification client is targeted at three classes of users:

- existing Brainware customers that have remote workers
- business process outsourcing (BPO) organizations that leverage keyers in developing countries
- OEM partners looking to embed a broader offering of services in their cloud-based applications.

The support for double-byte character recognition specifically applies to Chinese, Japanese, Korean and Thai characters. “This was driven by customers doing business with Asian manufacturers,” said Kaplan. “The OCR vendors we work with have finally reached a high enough accuracy level with their CJKT engines that we feel comfortable deploying them in *Distiller* for content-based data extraction.”

Brainware's new browser-based *Gateway* module is designed to reduce the overhead associated with answering vendors' questions about invoices. “Basically, after an invoice is posted, *Gateway* can be used to automatically fire off an e-mail to the vendor that issued it; the e-mail includes a secure link into the system,” said Kaplan. “*Gateway* is connected with the user's ERP system and provides a status report based on information from that system. Since most organizations devote resources to answering calls about invoice status, the ROI opportunity is obvious.”

Rapid deployment for A/P

With the release of V5, Brainware is also heavily promoting *Distiller's* rapid deployment capabilities. “We can get a customer up and running on our accounts payable package in as little as 48 hours,” said Kaplan. “This includes our standard integration into ERP packages from vendors like **SAP**, Oracle, **J.D. Edwards**, **PeopleSoft**, and **Lawson**.”

“Our current A/P package is based on the experience we've gained from doing hundreds of installations. It has a very robust set of capabilities for capturing header and line item detail, as well as doing validation. It's not uncommon for us to successfully get 90-plus percent field level extraction on a set of invoices we've never seen before. Most services related to our implementations have to do with exception handling.”

Hyland latest ECM partner

Brainware also recently announced a partnership with imaging-centric ECM ISV **Hyland Software**. The partnership is based on integration between *Distiller* and Hyland's *OnBase* software to create

“end-to-end invoice processing solutions for A/P departments.”

“We are seeing *OnBase* in an increasing number of accounts,” said Kaplan. “While Hyland has had partnerships with various other vendors in the data capture space, they recognize that working with us will enable them to go more aggressively after larger, more complex accounts. This partnership involves mutual referrals, and we recently worked with Hyland and one of its reseller partners on an implementation [at a large global manufacturer].”

When we asked Kaplan what capabilities differentiate Brainware from Hyland's other capture partners, he cited the company's fuzzy matching capabilities as one example. “Our background as a developer of search technology enables us to reconcile invoice items with P.O. items—even if it's not an exact match,” he said. “It's not unusual for a company to issue a P.O. for ‘a case of black felt-tip pens’ and the invoice to come for something like a ‘crate of thin markers.’ Even though both represent the same thing, if you are looking for an exact match, you're not going to get it. *Distiller's* ability to reconcile fuzzy matches enables us to increase our rate of straight-through processing.”

For more information: <http://www.brainware.com/>;
<http://www.brainware.com/news.php>;
<http://tinyurl.com/BrHyland>

Vendor Offers Scalable, E-Discovery-Centric Search

I've written quite a bit in *DIR* over the past couple years on the potential for ECM vendors in the emerging market for e-discovery systems. However, it's also important to realize that an ECM system is not an e-discovery system by itself. While an ECM system can be used to store and organize records that may be called upon as part of an e-discovery process, there are several steps related to collecting and reviewing legal content that many ECM systems do not optimally address.

Historically, the market for electronic review has been dominated by vendors like Concordance and Summation, while PC Docs and iManage/Interwoven have done most of the back-end electronic document management related to collaboration and collection for legal cases. However, as electronically generated information has surpassed paper as the main target of legal discovery, new software is being developed specifically to meet the demands of e-discovery.

In the past, I've featured **Anacomp's** *Caselogistix*

application, which offers innovative review techniques designed to take better advantage of modern technologies [see *DIR* 7/3/09]. Recently Anacomp, which has a long history as an imaging service bureau and offers hosted document repository services, announced a partnership with **Digital Reef** to introduce advanced collection capabilities into its e-discovery platform.

Digital Reef is a Massachusetts-based ISV that has developed a "similarity engine" designed to quickly crawl large bodies of information and "determine which files are similar in content, metadata, and structure." "Our search technology is designed specially to address the e-discovery market," said Steve Akers, the founder and CTO of Digital Reef. "We get the question all the time, if I have enterprise search, why do I need to re-index my files for e-discovery? Enterprise search engines are built for e-commerce. This means they have different sets of synonyms and indexing techniques that aren't optimized for e-discovery.

"Our engine searches both full-text and meta data, as well as information like who modified a file, when it was modified, where it's residing, and the file-level permissions associated with it. All these factors go into legal defensibility."

Akers has experience managing large blocks of data. In the late 1990s he founded a company that produced switches for delivering data over IP networks. In 2000, he sold it to **Lucent** where he worked before launching Digital Reef a couple years ago.

With Digital Reef, Akers set out to address legal

issues associated with anything that lies "outside of databases." "Many people call this unstructured information; traditional search engines can help you locate it based on keywords, but that really doesn't help you when it comes to e-discovery," he said. "We also heard feedback that people couldn't find enterprise-class tools that could break down their data into pieces their review systems could digest.

"We built our technology to go after multi-terabyte collections, even hundreds of terabytes. After all, most people now have a hundred gigabytes on their laptops. You can imagine the amount of data in any reasonable size company with 1,000 employees."

Akers stressed that the stuff Digital Reef is doing is now mostly being done manually. "Legal departments and law firms have humans looking at thousands of documents to decide which belong together based on similarity of content," he said. "They have to read and tag these documents manually. We can save tens of thousands of man hours by being able to intelligently classify them. This works very well in applications like early case assessment."

Akers said that Digital Reef has built connectors that enable it to access information in most major ECM systems. "We are also looking to form partnerships with any vendors that deal with a lot of unstructured information," he said. "We offer our technology as both software and as a hosted service. We are definitely open to OEM agreements. In addition to e-discovery, we can help with areas like storage management, compliance, and risk auditing."

For more information: <http://www.digitalreefinc.com/>

Subscription Order Form for RMG Enterprises, Inc.

4003 Wood Street • Erie, PA 16509 • Phone (814) 218-6017 • e-mail: ralphg@documentimagingreport.com

Document Imaging Report

Business Trends on Converting Paper Processes to Electronic Format

SCAN: The DATA CAPTURE Report

Premier Management & Marketing Newsletter of Automatic Data Capture

1 year (24 issues)

- ☐ electronic copy @\$597
☐ paper copy @\$670

- ☐ electronic copy @\$597
☐ paper copy @\$670

OUR GUARANTEE TO YOU
 If you are not completely satisfied, we will refund your subscription cost for all remaining unserved issues.

Name _____ Title _____

Company _____

Street _____

City _____ State _____ Zip _____

Phone (____) _____ Fax (____) _____

E-Mail _____

Please ☐ enter / ☐ renew the following subscription.
 (Add \$33 on all orders outside of Canada and the United States.)

☐ **Payment Enclosed** (Remit to: RMG Enterprises, Inc., 4003 Wood Street, Erie, PA 16509)

☐ **Charge My Credit Card** (Charge will appear as RMG Enterprises.)

____ AmEx ____ Visa ____ MC ____ Discover _____ card number _____ expire date _____

☐ **Bill My Organization** (Purchase order # optional.) _____