

# 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>

September 14, 2012

## THIS JUST IN!

### ABBYY ACQUIRES CAPTURE RESELLER

**ABBYY USA** has acquired reseller partner **Digital Documents** (D-Docs). D-Docs is a document capture and ECM specialist with 250-300 customers nationally and just less than 20 employees. ABBYY, which develops document capture and recognition software and tools, plans to allow D-Docs to run autonomously, while leveraging D-Docs' knowledge and experience to help it continue to build out its VAR channel.

"We will treat D-Docs like any other VAR," said Dean Tang, president and CEO of ABBYY USA. "We want it to grow and to continue to carry other ISV's products. We just expect D-Docs to add one line to its marketing materials—an ABBYY company."

Digital Documents represents the first-ever acquisition for ABBYY USA. "We've been working with D-Docs for four years now," said Tang. "Historically, D-Docs has had a lot of success selling the Cardiff [now part of **HP's** Autonomy group] product line. It has also been a standout selling our FlexiCapture solution, and that is really what caught our attention."

"We studied the company and got to know the owner, Bruce Genger, at a personal level. He has a real entrepreneurial spirit that we thought would be a good fit for our company. We see good growth potential for D-Docs, and it has had a good first half of 2012, which mirrors the success that ABBYY has been having."

In addition to capture, D-Docs sells ECM offerings from **Hyland**, **Open Text** (Alchemy), and **EMC** (ApplicationXtender).

Digital Documents is especially strong in the

**CONTINUED ON PAGE 8**

## Bringing More Intelligence to Capture

GLEN COVE, NY—Bringing more intelligence to the capture process was one of the major themes that emerged at last week's **Harvey Spencer Associates'** (HSA) eighth annual Capture Conference held at the Glen Cove Mansion. Close to 100 attendees, mostly document capture hardware and software vendors, gathered to discuss the latest trends and gain insights into the future of what HSA measured as a \$2.5 billion software market in 2011.

A full day of speakers covered topics like semantic understanding, the transition from document capture to big data analytics, **SAP's** HANA appliance, security and governance related to cloud storage, mobile e-forms, and knowledge liquidity and intellectual property management. As usual, *DIR* Editor Ralph Gammon kicked off the conference with his predictions for the upcoming year. He was followed by HSA Principal Harvey Spencer, who reviewed market progress in 2011 and offered insights into where capture is headed in the future.

One of Spencer's observations that, "Today we deal with document capture, and tomorrow we deal with big data," seemed especially prescient. Spencer discussed expanding what he has traditionally defined as 'document capture' to embrace input from multiple channels, including photographs, voice recording, e-mail messages, video, Web sites, and social media. "However, there must be a business driver—a transactional input—that is where I draw the line," he stressed.

Spencer noted that document capture's underlying use of pattern recognition and business rules to understand and capture unstructured data can be applied across a variety of inputs. "One of the trends we are seeing is corporations dumping white collar jobs on consumers," he said. "For this trend to be effective, the processes have to change and become more consumer-friendly."

"Part of this involves adding more intelligence to capture. Data needs to be extracted in real time, so a user can be provided with immediate feedback and

asked for more information if necessary. For example, if a person uses a smartphone to submit a picture of an accident, it should naturally include information like who, when, and where [obtained from the smartphone system]. Pattern recognition could be applied to the image, which would enable an application to automatically provide feedback like, 'we need another image.'

Mark Hakime, chief architect, digital platform, **TD Bank Enterprise Association**, echoed Spencer's call for more intelligence in the capture process during his presentation on what end users want from vendors. "We want immediate feedback on our documents," Hakime stated. "We want to be able to OCR all the text and then run rules to check dates, signatures, addresses, and even information like length of employment to verify eligibility for a loan. We're looking at semantic technologies for that.

"We're also looking at semantics for areas like processing insurance claims. If a doctor writes that someone has a broken fibula, we want to be able to understand that they're talking about a broken leg. Capture is no longer just about digitizing documents and classifying them. It's about making the content on the document useful.

"Across our organization, we're dealing with a couple hundred million transactions. If we can shorten the time to process documents, it could be a huge savings."

Hakime concluded that TD Bank wants to apply capture to any data stream. "I think capture and semantic understanding should be folded into a single application," he said.

### **Semantics and predictive modeling**

This is the type of application that capture ISV **ABBYY** seems to be working on, as Aram Pakhchanian, VP, director of data capture products department, presented on this topic. Not that it is an easy task. "We have been developing semantic technology for 17 years, and we still don't have a product," he said.

Pakhchanian defined semantic understanding as "the ability to automatically extract entities, facts, and events from unstructured documents, along with their features and relationships between them. "It has to be based on some type of language understanding," he said.

Pakhchanian explained one case scenario of an intelligence agency being able to use semantic technology to extract not only people's names, but, also being able to figure out if two people were ever in the same place at the same time.

He listed one of the ultimate goals as eliminating any manual intervention in meta data extraction. "Human intervention naturally introduces some bias," he said.

Finally, Pakhchanian asked whether eventually meta data extraction would be removed from the capture process altogether. "Will it eventually end up as something being

## **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. 22, No. 16



**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 23x 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** © 2012 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.

done in the back-end, like data mining?"

So, what will users do with all this information provided to them by semantics. This was a question that Dr. Steven Schlosser, senior scientist at **NovoDynamics**, attempted to address. According to Schlosser, predictive modeling is the next big trend in data analysis.

"Today, the biggest uses of data are for historical trend analysis and standardized reporting," he said. "But two years down the road, those areas are projected to slip down the priority list behind such uses as data visualization and simulations and scenario development."

Schlosser described one use case in which software for an on-line shopping site automatically adjusts prices by analyzing data on customer behavior, competitive response, and other factors. "The Mackenzie Institute has projected that retail margins can be increased by 60% through optimizing data management and analysis," he said.

NovoDynamics has its roots in a research lab with expertise in image analysis, pattern recognition, and data mining. It has developed its own OCR/ICR and image classification and extraction software [see *DIR 10/8/10 & 4/6/12*]. Schlosser discussed not only analyzing textual data, but also photographs, videos, and medical images. "Some of the challenges we face as we attempt to move forward are that photographs, video, and medical images have high-value information but are difficult to analyze," he said. "Another area of input, social media, has a tremendous amount of content, but it's very noisy."

Schlosser concluded that the challenge for managing big data is going to be extracting the right data.

### **Next year's dates set**

This concept of data analytics converging with capture was just one of several hot topics that were discussed both online and offline at this year's HSA Capture Conference. We'll have more from the event in upcoming issues. Once again, I'd like to thank everyone at HSA for their usual hospitality and say it was great catching up with those of you I saw last week. Next year's event has been scheduled for Sept. 11-12.

For more information:

<http://hsassoc.com/documentcapture/index.html>

Please don't forget to visit our *Document Imaging Talk* blog at <http://documentimagingreport.blogspot.com/>. Recent posts discuss **Kofax's** Q4 results and review Brad Paxton's book on forms processing. Also, for news updates, follow us on Twitter @DIREditor.

## **New Software Addresses Remote Scanner Management**

Way back when, when document capture was strictly a centralized back-office operation, managing a scanner was a fairly straightforward process. All of an organization's scanners were operated in the same area and an administrator could oversee and manage them directly. If a scanner was underperforming, the supervisor could walk over and take a look. If a driver needed to be updated, an administrator could physically load a CD into the PC attached to the scanner.

However, as distributed capture implementations have emerged over the past 10 years, the dynamics of scanner management have changed. Now organizations often find themselves managing scanners spread out in 10s and even 100s of remote locations. The volume at many of these locations does not warrant a scanning expert on site, and, in many cases, the scanning is being done by a knowledge worker whose primary expertise lies elsewhere.

Thus, our industry now faces the challenge of remote scan management. "We have a lot of customers purchasing a large number of scanners, and the IT staff wants to be able to manage them remotely," said Will Hebert, product manager for **Kodak's** new Asset Management Software (KAMS). "They want to be able to do things like track throughput, manage scanner IDs, and handle maintenance and updates centrally. And, quite frankly, for any USB peripheral, including scanners, there are not a lot of tools available for doing that type of stuff.

"We've found a lot of our customers were MacGyvering solutions together, which includes using scripts and log-ins just to remotely install drivers. But, they really weren't accomplishing their goal of proactive management of their scanners."

Hebert said Kodak conducted an informal study of its customers. They asked the scanner vendor for 10 deliverables in scanner management software. These included data collection, reporting tools, consumables tracking, condition alerts, easy access to log files, scanner identification and registration, the ability to group scanners by several criteria, and scheduled deployments of drivers and updates.

"Before we came up with KAMS, we wondered if there was any way this sort of functionality could be delivered with some of the tools already available on the market. We looked at PC lifecycle management software, and while we found it was very good for

managing PCs, it doesn't work so well for peripherals. On top of that, it's expensive.

"We looked at other types of application monitoring software in the ECM market, from vendors like **Reveille** and **Mavro**, but it really doesn't have the appropriate level of device control that IT staffs are looking for.

"Finally, we looked at what we consider the Holy Grail of asset management software—**HP's** Web Jetadmin. It's free and it handles other vendors' MFPs. The problem is that it utilizes SNMP [simple network management protocol], which requires a MIB [management information base]. And while the MIB for the printer market is well-defined, there are no MIBs for scanners. It took the printer industry 10 years to work out their MIBs, and the standard definition of scanners is not nearly at the same level of maturity."

### **How KAMS works**

According to the press release, Kodak's new Asset Management Software (KAMS) is a thin-client application that "allows users to register, install, and configure devices. [It] also enables technical staff to remotely monitor and troubleshoot scanners in the enterprise. Users and administrators can receive alerts and set thresholds for paper feeding and consumables usage. The software can also manage, store, and secure scanner profiles and user settings. It supports multiple login accounts with different levels of privileges for secure operations."

"The server piece of KAMS takes advantages of Web services," said Hebert. "This way it can utilize standard Web ports in an organization's firewalls. Administrators can access it through several types of Web browsers. The only requirement is that their device must be able to run Silverlight, which means KAMS can't be accessed on an iPad yet.

"On the PC attached to each scanner, users install a collector. The standard way to do that is through a PC lifecycle management tool or a silent install script."

KAMS stores information such as model, serial number, and the scan count of a device. Scanners can be configured in groups and sub-groups with access rights restricted to specific administrators. "This is important to organizations with an especially large number of scanners," said Beth Schmidt, training manager at Kodak. "KAMS can manage multiple accounts with different levels of privileges granted to each account. You can control what assets a particular administrator can see, as well as whether they can configure certain aspects of the system."

Drivers, profiles, and firmware can all be deployed and updated through KAMS. "This includes profiles for our Smart Touch capture application," said Hebert.

A SQL database is used to store all the scanner attributes as well as information like when consumables should be changed, what type of throughput each scanner is getting, how many paper jams are occurring. Reports can be created through off-the-shelf tools like Excel and Crystal Reports. "The information in the SQL database can be sliced and diced in multiple ways," said Schmidt. "If you have knowledge of reporting tools, it's limitless what you can come up with."

Text and e-mail alerts can also be designed to notify an administrator when scanning levels drop below a certain threshold. "The threshold for a group of scanners might be determined by the model, as well as the type of documents being scanned," said Hebert. "Expected throughput, and therefore threshold, might be lower for scanners capturing difficult documents like bills of lading vs. scanners capturing something simpler and more straightforward."

KAMS does not need to get feedback from devices in real time but it does allow the communication interval between the server and each device to be configured. "The default is once a day, but for someone running several thousand devices, that might not be their cup of tea," said Hebert. "They might want to scale up to once an hour or once every few minutes."

The first version of KAMS is designed to work with distributed [workgroup and departmental] Kodak models: the i900 series, the i2000 series, and the i1200/i1300 Plus series. "Distributed scanners represent the greatest number of devices in the field and distributed scanning implementations are where we are hearing the most requests from customers for this type of software," said Hebert. "With the next version, we will add support for some production models and widen the coverage within our portfolio. That version is not too far off."

Hebert said that supporting other vendors' products is more of a challenge. "It has to do with the concept I discussed before recording MIBs," he said. "Scanners just don't have a mature enough architecture to enable us to manage our competitors' devices through KAMS—without us basically reverse engineering their drivers, and I don't think they'd take that too well."

Hebert indicated that even Kodak's network scanners, for which there already is a centralized

management tool [see *DIR* 2/19/10], can't be managed through KAMS either. "KAMS is a completely different product than our network scanner management tool, because the architectures of the hardware are so different," he said. "A network scanner is a scanner, plus a computer, plus software applications all rolled into one. The communication path you have to take to manage a network scanner is way different. Plus, there are some major differences in the use cases."

KAMS is available now, for a U.S. list price of \$3,500, which enables an unlimited number of units to be monitored. "We expected the sweet spot to be organizations with between 50 and 100 scanners," said Hebert. "Initially, we're finding it to be a little higher—around a couple hundred."

For more information:

<http://www.documentimagingreport.com/index.php?id=2333>;  
<http://tinyurl.com/KAMSpage>

## Document Compression Sales Growing

Advanced document compression used to be one of our favorite topics in *DIR*. But, as storage costs have continued to drop, and PDF has become the standard alternative to TIFF Group 4s for storing document images, it seems like compression has become less of an issue. Carsten Heiermann, the CEO of German compression and document capture ISV **LuraTech** agrees.

"For a long time, our pitch for PDF Compressor was that it helped create faster download and retrieval times," Heiermann told *DIR*. "Now, we take the approach that, of course, people want to have documents that can be stored efficiently and downloaded quickly.

"We are finding more of our customers making decisions based on their desire for compliance and adherence with the PDF/A standard [PDF/A files can be created through PDF Compressor]. They are also looking for software that can create PDFs with a reliable level of quality. The PDF software market can be a nightmare—there are so many bad PDFs out there that people can't view.

"We have developed a strong reputation from our work with **ISO** on PDF standards, and as a founding member of the **PDF Association** [which grew out of the PDF/A Competence Center, see *DIR* 10/22/10]."

According to Heiermann, compliance with PDF/A (an archiving standard for PDF documents, see *DIR*

5/18/07), is still more prevalent in Europe, especially in Germany, but it has certainly found some niches in North America. "In Germany, and some of the surrounding areas like the Netherlands, utilizing PDF/A is almost a no-brainer. If someone is scanning files as part of a document management project, output to PDF/A is almost always included. Adoption is past the tipping point and really rolling.

"In the U.S., you get a lot more questions like, 'Do we have to use it?' and 'What is it going to save me?' That said, there are some industries, and I just gave a talk for the **Nuclear Information and Records Management Association**, that are really into PDF/A. It's from relationships like we have with NIRMA that we've been able to win some large deals recently."

LuraTech prefers to work through a VAR channel, and Heiermann said many of LuraTech's PDF Compressor software sales include a healthy level of professional services. To better address these large deals, LuraTech recently came out with a new version of its software, PDF Compressor 6.0, which is faster, more accurate, and more flexible.

"From a feature perspective, I think we already had a pretty complete product," said Heiermann. "That said, the new version is a fairly major redesign to improve our performance and throughput. That's important to our customer base as many of them move from trying PDF/A on a limited scope, to a bank, for example, now using our software to convert all its credit files to PDF/As.

PDF Compressor 6.0 has integrated **ABBYY** FineReader 10, which improves its performance on "noisy, low-contrast, or distorted scanned documents." There is also a new OCR "Fast Mode." "We've also improved our compression, including the quality, so there are even fewer errors," said Heiermann.

LuraTech's success in high-volume PDF conversions has led it to a secondary market—selling its DocYard software to service bureaus for managing production scanning projects. "DocYard tracks everything from inventories before scanning, to how long document prep is taking, to at what point in production a batch is, all the way through to when electronic images are delivered to an ECM system and paper documents are either stored or shredded," said Heiermann. "We have one service bureau using it to manage scanning of more than 50 million pages per month."

"We've been able to succeed in the service bureau market because of our experience in high-volume capture environments. PDF Compressor is known not only for creating quality PDFs, but for

consistently working successfully on projects with several hundred thousand pages. Once PDF Compressor is set up, there is almost no maintenance. There should be very few errors and no crashing of the application. We stake our reputation on reliability and quality.”

For more information: <http://www.luratech.com/>;  
<http://tinyurl.com/PDFCompressor6>

### **CVision targeting VARs**

Like LuraTech, **CVision** has a product called PdfCompressor (one word), used for creating optimized PDF and PDF/A files. CVision’s Maestro Recognition Server is probably more directly competitive on the type of high-end projects that Heiermann discussed. Like LuraTech (which Heiermann reported saw 50% growth in the first half of 2012), CVision has also discussed some strong growth numbers.

However, CVision differs from LuraTech in that it feels the market for compression technology is anything but played out. In fact, CVision is currently planning a new VAR program with its compression technology as the centerpiece. “We think there is a tremendous opportunity for VARs for optimizing PDFs as end users move more images to the cloud and want to download them onto mobile devices,” said Scott Harvey, director of channel sales for CVision.

“First of all, when you use cloud storage, you typically pay per gigabyte. So, utilizing compression technology to reduce file sizes is a money saving strategy. In addition, as more users seek to download documents onto their mobile devices, smaller files eat up less bandwidth.”

Added Chris Koulouris, director of marketing for CVision, “We think VARs like to talk about the cloud and mobile, and our technology gives them a reason to discuss those topics.”

Harvey thinks compression technology could be attractive to the channel because it’s an additional product that doesn’t displace anything. “It’s really a compliment to their ECM and capture products,” he said. “We’re projecting an average deal size of around \$10,000, which includes \$8,000 in software and the rest in professional services. While that’s a decent sized sale, it shouldn’t be large enough to raise any red flags for end users.”

In many cases, Harvey noted, the savings on storage alone may pay for the compression software. “When you start looking at opportunities, they include not only backfile and day forward paper scanning, but also backfile conversion of electronic

documents,” he said. “Even if users are converting black-and-white TIFFs to PDFs, we often see a size reduction of 5-to-one.

“We have one reseller working with a hospital where there is a day-forward mandate to scan all paper and save the images as PDFs. That hospital already has billions of electronic documents that there is no requirement to convert. However, if they do convert those backfiles using our compression, it will more than pay for the cost of our software by the server space being freed up.”

Like LuraTech, CVision has a secondary product, Trapeze for Invoices [see *DIR* 4/6/12], which it plans to eventually bring to market through the channel. “Right now, I think our compression products are more channel-ready than Trapeze,” said Harvey.

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

## **Snowbound Upgrades AJAX Viewer**

Last time we talked with **Snowbound Software** CEO Simon Wieczner he cited increasing adoption of the company’s AJAX viewing software as helping to drive strong revenue growth [see *DIR* 5/4/12]. To better address the enterprise implementations where VirtualViewer AJAX is increasingly being used, Snowbound recently released a new version, 3.0, that features multiple improvements. These include a new user interface, improved annotation and page manipulation, and expanded international language capabilities.

“This is a fairly major re-write,” explained Wieczner. “There was a lot of stuff that worked well enough before, but that we decided to redo to make the product more enterprise worthy.”

VirtualViewer AJAX is a zero-footprint viewer designed to be accessed through a browser. “A lot of our customers really want to get away from depending on thick clients,” said Wieczner. “They want to do everything in an AJAX viewer that they can with a traditional thick client viewer.”

This includes page manipulation. “Users can now delete, copy, and insert pages, and create new versions of documents directly within their viewers,” said Jody Spencer, director of marketing for Snowbound. “They can copy and paste from any type of document and utilize three documents open in different tabs to create a fourth document that can be saved as a PDF, JPEG, or TIFF.”

VirtualViewer AJAX 3.0 also offers more

customizable annotations. “We’ve given more much flexibility to the end user,” said Wiczner. “They can now control all kinds of things, like the color of the background and foreground. Of course, an administrator can also limit the flexibility.

“We’ve also changed our API, which governs how annotations can be accessed when our viewer is integrated with third-party products. You can now search annotations from third-party applications, for example.”

Snowbound has pre-built integrations for VirtualViewer with ECM applications from vendors like **IBM** (FileNet), **EMC** (Documentum), **Open Text** (Livelink), and **Microsoft** (SharePoint).

To address its increasing number of international deployments, Snowbound has added the ability to make annotations, as well as view menus in multiple languages. “Basically, our software is able to detect the language of the OS client and will match its display to that language,” said Wiczner.

Wiczner concluded that Snowbound’s use of true AJAX technology gives it an advantage in the browser-based viewer market. “You hear a lot about AJAX viewers, but if you look under the hood, you see Flash or HTML 5,” he said. “This means these applications won’t display the same on certain devices or with some older browser versions. When you’re dealing with some of the large organizations that we deal with, being able to display cross platform can be a big deal.”

For more information: <http://www.snowbound.com/>; <http://www.documentimagingreport.com/index.php?id=2325>

## Automated Data Capture Still Underutilized

The numbers are fairly sobering. Like a splash of cold water to the face after a late night out. Yes, despite our tireless promotion of automated data capture technology leveraging OCR and ICR, only 33% of professionals surveyed in a recent **AIIM** study said they apply OCR to forms to capture machine-print data. This is despite the fact that 88% of the respondents are scanning forms. [Source: AIIM White Paper: *Forms Processing - user experiences of text and handwriting recognition (OCR/ICR)*]

In another recently published AIIM study, although more than 90% of respondents said they were doing scanning, less than 20% said they were doing “full data capture to multiple business processes,” while more than 40% were doing manual indexing for

routing/archiving. [Source: AIIM White Paper: *Distributed and Mobile Capture-moving the process closer to the customer*].

In both these surveys respondents were members of the AIIM community, so presumably they have a higher than average understanding of ECM technologies, including capture. You can imagine that the adoption rate of data capture technology in the world at large is even lower.

The forms processing white paper was sponsored by **Parascript**, which develops OCR/ICR and cursive writing recognition technology. “What precipitated us to do the survey was that we’d be telling potential customers what we did and they’d say, ‘that works?’” said Dwayne Ritchie, VP, sales and marketing, at Parascript. “The survey supports our belief that people do not have an understanding of the products available in the market.

“And discussions we’ve had with IT groups and the analyst community confirm that they see about the same rate of adoption [as in the survey] for automated recognition.”

Added Don Dew, senior marketing manager at Parascript, “When we tell people what we do, we get a lot of surprised looks. Part of what we are up against is poor education about the capabilities of automated recognition.”

As you can imagine, adoption rates of handprint and cursive recognition were even lower than the use of OCR for machine print. According to the AIIM forms processing white paper “12% use ICR to recognize hand-print constrained field entries. 6% use ICR to recognize hand-written script and free-form entries.” “We really didn’t know what to expect regarding adoption of hand-print and cursive recognition, but we really weren’t surprised,” said Dew. “Basically, the survey results serve as a valuable touchpoint to validate that there is a tremendous opportunity in the market around forms processing.”

### Upgraded SDK

To address this opportunity, Parascript has introduced a new version of its FormXtra SDK for automated data capture. FormXtra 5.0 introduces semi-structure document capture capabilities as well as a new API that makes available some 1,400 discreet commands.

“Our previous version did a great job on structured forms,” said Dew. “We’ve added capabilities like word-spotting and the ability to define dynamic fields like the ‘total’ on an invoice by setting up rules. We’ve also introduced the ability to capture data from tables, which can be of varying lengths.”

Dew acknowledged that semi-structured forms capture is not necessarily revolutionary. "However, what differentiates FormXtra is its ability to capture a wide variety of data types—machine- and hand-print, as well as cursive, in a single application," he said.

Dew said the new API configuration increases the versatility of FormXtra. "In the past, you'd basically define the form up-front and then pass it on for recognition and data extraction," he said. "While you can still do that, you can also call up more functionality on an individual basis. For example, if you want to do image processing and then review an image manually before submitting it to the next step in a workflow, such as redaction, you can do that.

"We previously only had about 50 discreet commands. We've also moved to a .NET API. Our plan is to continue to broaden the capabilities of our API. In the future, you can expect to see even more ability to call on just the services you need."

Parascript has also introduced new field-level voting into FormXtra, including the ability to utilize ABBYY's OCR for machine-print recognition.

"Historically users have needed two or three different recognition SDKs to accomplish what they can now accomplish with FormXtra," concluded Dew. "Being able to look at a total document with a single application addresses one of the shortcomings of current products on the market and should help drive more adoption of recognition technologies."

For more information: <http://www.parascript.com/>;  
<http://tinyurl.com/FormXtra5>;  
<http://www.parascript.com/aiim-and-parascript-study>

## ABBYY D-DOCS, FROM PAGE 1

healthcare and education markets. Its marquee customers include **Healthways, Stanford University Hospital and Medical Centers, American Medical Response, the Duke University Dept. of Surgery, Johns Hopkins University Center for Talented Youth, UCLA, and Emory University.** D-Docs also sells into the government and financial services markets.

"We see the healthcare market as a huge opportunity for us," said Tang. "But, we need to break into it through partners. D-Docs has developed the experience, know-how, and best practices needed to be successful in that market, and we think we can leverage all that to help our other partners be successful as well."

Tang stressed that ABBYY's lead program for VARs will be unaffected by the acquisition. "Our current focus is going step-by-step to make sure we integrate the acquisition successfully," said Tang. "If we utilize the correct methodology, all our partners should benefit."

"We will work with D-Docs to figure out the best practices for going to market through a channel. [Similar to the way ABBYY has always utilized its shrink-wrapped OCR application to help it study the market for its recognition SDK.] We then plan to apply those best practices across our entire reseller channel."

For more information:  
[http://www.abbyy.com/News/ABBYY\\_USA\\_Acquires\\_Digital\\_Documents/](http://www.abbyy.com/News/ABBYY_USA_Acquires_Digital_Documents/);  
<http://www.d-docs.com/index.htm>

### Subscription Order Form for RMG Enterprises, Inc.

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