Document Imaging Report

Business Trends on Converting Paper Processes to Electronic Format

June 21, 2013

COUPLE OF RECENT LARGE CAPTURE WINS

Congratulations to **BancTec** and **ReadSoft** on a pair of recently announced huge capture deals. BancTec, a Dallas-based transaction capture hardware, software, and outsourcing specialist, recently installed one of the highest volume document capture systems in the world at a Japanese BPO center. ReadSoft, meanwhile, announced the first U.S. enterprise sale of its foxray XBOUND capture platform.

BancTec, working in conjunction with Japanese partner **J-SCube**, installed a system for capturing 3.5 million documents daily at the **SG Holdings' Sagawa Express** BPO Center. Sagawa Express is one of Japan's leading logistics companies. The system has been in operation since September and includes 30 of BancTec's very high-volume IntelliScan XDS scanners.

"BancTec's IntelliScan solution replaces Sagawa's legacy system by offering better overall OCR recognition, improved processing efficiency, and the ability to process a wide range of paper types and sizes," said Kazunori Kamata, division director of the project division at J-SCube, in a press release.

The BPO center is located in Koriyama City in the Fukushima Prefecture, which was a region affected by a nuclear disaster following the 2011 earthquake and tsunami that hit Japan. "We are very proud of this installation, the largest in the world for BancTec, and possibly the largest scanner center anywhere," said Mark D. Fairchild, SVP and CTO at BancTec. "A key reason for locating the BPO center in Fukushima was to revitalize the area and provide employment. So we were honored to

ISV Targeting SMB with New Cloud DM App

At last year's **Kodak** Global Directions Conference, noted industry analyst Rai Wasner put on a fascinating presentation on the potential of consumer apps like **Box.com**, **Dropbox**, **Google** Drive, and a host of others that he sees being used as ad hoc document management applications. It was Wasner's vision that these free and inexpensive services provided enough of what most people needed, in terms of file management, that they were going to jeopardize the future of traditional ECM.

Obviously, this type of legacy ECM apocalypse is at least a few years off—and probably a bit exaggerated in the same way the SharePoint-driven ECM apocalypse has been somewhat overhyped. That said, Wasner isn't the only one who has observed that these user-friendly, inexpensive file sharing repositories are being utilized for entry-level ECM. (In fact, Shad White, the entrepreneurial founder of **CloudPWR**, was at Wasner's presentation and a few months later announced that his company had completed a cloud-based procurement system for the **State of Washington** that leverages Box.com's file system.) But, much like scanner vendors view MFPs as offering great technology for introducing businesses to the power of capture, some ISVs see the adoption of on-line collaborative applications as an opportunity for upselling more powerful DM.

etfile is one such ISV. The Westborough, MA-based ISV is planning a July 4 launch of OfficeScope—a hosted imaging and document management application targeting SMBs. "Dropbox and Box.com are great for file sharing and collaboration," said Jon Lincoln, business development manager at etfile. "But, not every document needs to be shared. Law firms, for example, want almost nobody to see most of their documents."

Lincoln highlighted five major differences between OfficeScope and popular file sharing platforms. "Our application offers security and analytics, intuitive searching, streamlined paper filing, mobile scanning, and custom integration to Salesforce.com and other

cloud-based business applications," he said. "These are all features required by businesses. Products like Google Drive were launched as consumer-based applications and are now being used by businesses. But, they are not geared for business-class document management."

Deep DM roots

OfficeScope may be a new product, but etfile is certainly not new to the document imaging market. The company started out as an Alchemy reseller in the mid-1990s and eventually developed its own insurance focused imaging and document management application. etfile's client/server software is now installed at more than 1,000 locations and has more than 30,000 users.

etfile has historically sold its software direct to end users. "Our biggest success is in the insurance market, where we have integrated our software with a number of agency management systems," said Lincoln. "But our software was never built to solely target insurance. In recent years, we have expanded outside the insurance market into areas like property management and education.

"But, insurance continues to be our biggest market because that is where we are the most established and have built a name for ourselves. We decided the best way to really expand into other markets was through resellers. We felt we could rely on them to supply their vertical knowledge. We thought a cloud-based application like OfficeScope might be more attractive to resellers and help differentiate us."

OfficeScope is being marketed in three separate editions, with the Professional version starting at $18 per month per user. A Business edition lists for $27 per month per user, and a Business Plus edition is available for a negotiable price. Users can also have a free trial.

"OfficeScope is really geared toward SMBs who are using Dropbox or Box.com because many times they don't realize there is anything else available," said Lincoln. "One way we are going to market is through a self-service model. We are working on some ways to incorporate resellers in this model—like giving them codes for end users to enter when they purchase seats."

OfficeScope is hosted on Windows Azure and will typically be deployed in a multi-tenant fashion. "However, if a Business Plus user wants a private cloud, we can set that up too," said Lincoln.

Search, scanning, key features

Lincoln showed us some of the product's differentiating features. One is its ability to search by contacts. "The user doesn't need a key word or a file folder, they just need a contact name," he said. "They can then use document filters to drill down to the specific file they are looking for.

"This is a great feature for users that deal with a lot of customers. We originally developed it for insurance agents,

but it can also be utilized by accountants, lawyers, and school districts—whose customers are the students. The contact search goes across all folders and cabinets and returns to the user a list of all documents they have secure access to."

OfficeScope's integration with **Salesforce.com** enables users to import all their Salesforce contacts to help with file indexing. OfficeScope is also integrated with Outlook and Office. The product features check-in and check-out and versioning capabilities as well.

OfficeScope brings to bear the scanning technology that is an important part of etfile's heritage. "Our initial product was positioned as a scanning solution to help companies go paperless," said Lincoln. "Over the years, we expanded to a full document management solution with e-mail integration and workflow. Not everything is being brought over into OfficeScope, at least not right away."

The scanning certainly is. OfficeScope features a zero-footprint client that can connect to TWAIN drivers. "We offer some intuitive ways of scanning paper that collaborative platforms just don't have," Lincoln said. "This includes being able to index through bar-coded cover sheets. We also have zonal OCR that we can set up for Business Plus users."

etfile is a long-time **Canon** scanner reseller and is looking at integrating OfficeScope with Canon's ScanFront 300 network scanner. "We think there is the possibility to sell OfficeScope as part of a bundle with hardware products like the ScanFront," said Lincoln. "We are also talking to MFP manufacturers about integrating with their Web-based interfaces."

OfficeScope has analytics capabilities enabling users to create custom reports. "Users can run reports on activities and users," Lincoln said. "For example, we had one customer go to court with records on document access that our system produced to prove that one of its ex-employees had stolen information. You can't do that with Dropbox."

OfficeScope will include a cross-format document viewer, which was one of several features being finalized when we spoke with Lincoln last week. An app to enable document access from mobile devices is also in the works.

"We're not going to have everything in the product right away when we launch it," Lincoln said. "But, at some point you have to make a commitment to go to market and start getting feedback from your customers. We plan to continuously introduce new features, some of which are already almost completed."

One feature planned for a little further down the road is an approval-type workflow. "With OfficeScope, we are really targeting the SMB, where we don't think the full-blown workflow that we have in our client/server product will be needed," said Lincoln.

etfile will continue to support and develop its client/server product. "That product will continue to be targeted at the insurance market, where it is integrated with several back-end systems," he said. "We are looking to integrate OfficeScope with more CRM systems like SugarCRM and possibly ACT."

VARs will be able to sell either the client/server or the cloud version of etfile's software. "We will also be replacing the Web-based interface that we currently offer with our legacy product with the OfficeScope technology," said Lincoln. "This will help us better address demand for hybrid environments."

Lincoln concluded that etfile is both nervous and excited about the upcoming Independence Day launch of OfficeScope, which will carry the theme of freeing businesses from their paper. "Everyone recognizes that the software market is going to the cloud," he said. "The old software model of selling $30,000-$40,000 systems is being replaced by $800 subscription fees and some investment in training. "

For more information: http://www.officescope.com/ http://www.etfile.com/

Capture Start-Up Receives Series A Funding

The document capture market may be maturing, but that doesn't mean it still doesn't have some exciting potential. At least that seems to be the opinion of a group of three investors who recently put $2.4 million, through a round of Series A financing, into Berkeley, CA start-up **Captricity**. Captricity has developed a cloud-based service for data capture that relies on a combination of software and crowdsourcing.

Captricity is one of a new breed of capture ISVs that is relying on browser-based technology to expand adoption of data capture technology—which has historically been limited primarily to larger organizations and specialized service bureaus. "We've always had a mission of making organizations' data more accessible, whether it's on

paper, or fax, or PDFs," said Kuang Chen, founder and CEO of Captricity. "We've always wanted to be a one-stop shop for turning that unstructured data into 99% accurate structured information."

Captricity accomplishes this through a fairly innovative model. First of all, it charges a flat rate of \$.20 per page and anyone can sign up. High-volume discounts are available, but there is no minimum volume and the first 10 pages are free as a trial. The fee includes up to 20 fields per page, with \$.01 extra charged for each additional field.

To utilize the service, a customer basically uploads a sample form, follows a guided process for setting up a template, uploads the rest of their imaged forms, and receives back a data file. "One of our key goals is to make capture technology as easy to use as possible," said Chen. "That's one of our key points of differentiation.

Kuang Chen, founder and CEO Captricity

"That's important because we cater to customers of all different sizes. A customer who signs up for \$10 worth of digitization services should receive just as high quality of data as a larger customer. One use of our funding will be towards development that keeps making our software easier to use."

Who's investing?
The Series A funding was led by **The Social+Capital Partnership**, a Palo Alto-based group founded by former **Facebook** executive Chamath Palihapitiya. It lists its goal as "utilizing venture capital as a force to create value and change on a global scale." One reason that Captricity fits under the Social+Capital umbrella is that the company was founded out of Chen's graduate work dealing with helping third-world countries better process data. In fact, one of Captricity's early customers was a team of economists from **New York University** who were processing surveys filled out by Ghanaian farmers.

"As investors, we always seek to back companies with amazing ideas that make the world a better place," said Palihapitiya in a press release. "Captricity is making data work for organizations rather than the other way around, which is no easy task. The potential of data has been over-promised and under-delivered. What Captricity is doing shifts the balance in the right direction."

Palihapitiya has joined Chen as a member of Captricity's board.

In addition to Social+Capital, **Atlas Venture**, and **Knight Enterprise Fund** were also part of the Series A round. "Captricity is making a value proposition that you can keep paper workflow, but still adopt a useful and helpful digital system," said Elizabeth R. Miller of Knight Foundation. "What sets it apart? Data accuracy. Most of the existing automated solutions out there that are OCR-based and only get to 80% accuracy but you really need to be at 99%…[With Captricity] a mixture of human and machine intelligence gets you there without having to manually recheck everything. Before, you couldn't trust the results, and now you can."

Chen indicated that some funding could be used to invest in improved OCR. "Right now we use **Google's** Tesseract [open source] OCR engine and a host of other prediction and extraction algorithms we wrote ourselves," said Chen. "We use crowdsourcing labor to cover the delta between the results we get with our technology our goal of 99% accuracy.

"Utilizing a better OCR engine would be an obvious win because it could get us that much closer to our accuracy goal through technology alone and potentially decrease the crowdsourcing labor required. We haven't taken that step yet because we haven't felt we needed to."

Expanding into semi-structured forms
To date, Captricity has focused on processing structured forms like surveys, tax forms, tests, and applications. "We are also currently doing work on capturing documents like receipts, invoices, and business cards," said Chen. "But we haven't introduced those use cases into our product yet. We don't release anything to the public until we can unquestionably reach the 99% accuracy level."

Last September, Captricity released a mobile app for Apple iOS devices that enables users to capture forms with existing templates through smartphone and tablet cameras. "Like everything we do, we tried to make it as simple as possible—there is not a lot of processing work being done on the phone," said Chen. "It's our philosophy that the heavy lifting, like image processing, should be done on the server. That's where processing power is cheap so you can leverage a greater amount of capabilities."

The Series A funding brought Captricity's total money raised to \$4 million. Captricity had previously received \$1.6 million in seed funding. According to Chen, the seed money came from approximately a dozen investors—including the three participants in the Series A round.

For more information: http://bit.ly/captricitymobile; http://bit.ly/CaptricitySeriesA; http://captricity.com/

Market Definitions Changing

Is the line between document management and managed services blurring? True, there has always been a certain type of document management that has been classified as managed services, but this has mainly dealt with stuff like the mailroom— typically managing outgoing mail. That said, in recent years, there has been more crossover of the big mailroom management guys like **Xerox** and **Pitney Bowes Managed Services** into capture. On that note, I guess service bureaus could also have been considered managed services providers—but were typically more often categorized as outsourcers.

This brings us to an interesting point in the evolution of the market. As cloud software services start to ramp up—there seems to be some sort of intersection of traditional managed services, outsourcing, and software. Just look at the stories in this issue. For example, **Digitech** Director of Sales Sean Morris discusses how his company's hosted document management application is really considered by users to be an outsourced service. And David Bailey, co-founder of **DocuLex**, is all excited because his ISV has combined with three other companies to offer a broad suite of managed services, including server hosting and back-up and hosted document management.

That's not to mention what we're starting to see in the capture market, where ISVs have started to dip their toes into the pay-as-you-go model for cloud services. **Captricity**, which just secured a round of Series A funding, is a prime example of this. Captricity charges a flat fee per page captured. And long-time recognition technology ISV **Parascript** just released a new version of its FormXtra capture platform that includes a Web services deployment model and "a new, flexible pay-as-you-go subscription, and per-click based pricing formats."

Capture, once the realm of document outsourcers, is now being deployed as a managed service. My point is that old divisions of the market are now blurring and basically what we have now in the document imaging industry is some combination of managed outsourced software services. Like always, an end user can take a document, submit it to a capture process, and get back whatever they need from that document. But now instead of having to deploy that process in-house (very expensive), or outsource the process to a service provider (giving up control), the user can piece together their own managed services capture system, plugging in best-of-breed and –cost components to suit their needs.

Spencer Conference on horizon

This changing face of capture brings us to **Harvey Spencer Associates'** annual capture conference, being held this year the week after Labor Day, Sept. 11-12, at its usual location in the Glen Cove Mansion in Long Island, New York. As always, I expect to see several industry leaders there and look forward to stimulating conversations, interesting presentations, and generally good company.

And, as always, I will be presenting my annual industry predictions and reviewing my predictions from the previous year. However, Harvey has asked me to change my focus a bit to keep up with the changing themes of his conference.

If you've kept up with Harvey, you probably realize that he sees a convergence of document capture and the data it produces—with the whole concept of Big Data. And I probably shouldn't restrict that to document capture. Harvey, of course, prefers the more general term "capture," and, in fact, is currently working on a study to determine how voice recognition technology can be used to reduce Dodd-Frank Compliance Costs.

At Capture 2013, in addition to several traditional capture related topics, Spencer has added topics like Hadoop, Semantic Understanding, and Reliance on BPM at the Expense of Content. It should be a worthwhile event, and I look forward to seeing many of you there.

For more information:
http://www.documentimagingreport.com/index.php?id=2433

CVISION Launches Auto-Classifier for SharePoint

Microsoft SharePoint's increasing use as a document repository has certainly led to a number of capture solutions being designed to work with it. A lot of these solutions are focused on getting the right meta data attached to an image being filed in SharePoint. **CVISION** is taking a different approach with its new eFiler 2.0 software. The Queens, NY-based ISV's eFiler is an auto-classification application designed to automatically group files in the correct SharePoint folders.

"There are a couple scenarios where we see eFiler being utilized," said Shane Mitchum, a sales engineer at CVISION who gave us a demo of the product. "One is when users already have SharePoint libraries set up. eFiler can read the documents already in those libraries and then file

future images where they belong. The second is when users don't have any document management system and they want to set up one in SharePoint.

"In this second scenario, the users could set up their taxonomy or folder structure in SharePoint and then start scanning sample documents into it. They can also use eFiler to replicate a folder structure they already have set up in Explorer."

At the heart of eFiler is an engine that CVISION has deployed in the past in mailroom applications. It classifies documents based on text analysis, so full-text OCR is performed on every document processed. It learns by example and calculates confidence rates on each document it is attempting to classify. The confidence threshold before a document is sent for manual verification is adjustable. Afterwards, manually verified images can be added to the sample set.

"We suggest users have 20 documents to start to train the system," said Mitchum. "They can certainly use more, but they don't want hundreds. They can also start with as few as one."

Users can add key words to help with classifying unstructured documents. "A utility company might be trying to recognize which letters are complaints and maybe there are some words that typically appear in those types of documents," said Mitchum. "The keywords can be weighted."

The documents are output as full-text searchable PDFs, utilizing the technology that is also included in CVISON's PdfCompressor [*see DIR 2/1/13*].

In addition to filing documents in SharePoint, eFiler can save a copy somewhere else, like a file system. The images can also automatically be e-mailed. "Once a document is classified as an invoice, for example, it could automatically be e-mailed to someone in the A/P department," said Mitchum.

Mitchum noted that eFiler does not have to necessarily be used with SharePoint. "You could use it to classify documents for your Windows file system," he said. "SharePoint seemed like a nice easy starter repository that we should write a direct connection to. We plan to build connections to other repositories as demand dictates."

CVISION also has extraction technology that could eventually be incorporated in eFiler for capturing meta data [*see DIR 4/6/12*]. "Another option is that eFiler could be used to classify documents as different types, some of which the user may then apply automated extraction to," said Chris

Koulouris, director of marketing for CVISION. "For example, after a document is identified as an invoice, it could be sent to our Trapeze for Invoice Processing, or another vendor's invoice extraction software."

Koulouris stressed that one of CVISION's goals is to make eFiler accessible to all sizes of business. "We will have pricing and license options for any sized business," he said. "Usage could run from 5,000 pages per month up to 1 million. There is also a free trial version available for a limited number of documents."

And although the product will be sold through the channel, Koulouris expects many organizations will be able to install it themselves—which has historically been the case with PdfCompressor. "We are targeting the SMB, so we recognize that in many cases our customers will have smaller, less sophisticated IT departments and fewer resources to dedicate to a deployment," said Koulouris. "But, we still anticipate most customers will be up and running live within a day or two of purchasing eFiler."

For more information:
http://www.cvisiontech.com/products/general/efiler.html?lang=esp

DocuLex, Others Combine in Managed Services Play

Turns out **DocuLex's** partnership with its software hosting provider was just the first step of a larger plan. The Winter Haven, FL-based ISV recently announced it has combined with three other businesses to offer a range of managed services from e-mail encryption to data hosting and back-up. The new company is called **Protected Trust LLC** and all four entities will continue with their current businesses as well as try and combine with their new sister companies to address larger customer needs.

"We are trying to move our cheese a little bit from solely focusing on the document management market," said David Bailey, co-founder of DocuLex and now a VP of business development for Protected Trust. "A lot of the resellers that we work with are already starting to move more deeply into managed services.

"We've brought together four businesses that all offer some sort of managed services. What we have in common is that we all have what I call 'boutique'-type offerings. Our customers want privacy and security, and they want to know where their data is. They are paying a little more for the personalized

service that we provide."

The data and server hosting is provided by **Inland Fiber & Data**, which also offers back-up and recovery services. DocuLex sells capture and document management software that can be installed on-site or licensed through a hosted model. A company called **Elephant Outlook** hosts businesses Microsoft Exchange servers. And Protected Trust's legacy is in secure e-mail encryption and exchange.

All four companies are located in the same office building in Winter Haven and also share an investor. "All the services really support each other," said Bailey. "For example, there is a doctor's office nearby with three surgeons. We went in there to sell them a secure e-mail service, because they have to send sensitive patient records back and forth with the hospital all the time. Then we noticed their server was three-feet from the backdoor and susceptible to potential water damage, so we started talking to them about moving to a hosted server. In addition, we discussed moving them to our Exchange e-mail server from a Yahoo system so they'd be more HIPAA compliant.

"Protected Trust is a strong name that already had a trademark, so we went with that. But, we really see opportunities for all four businesses to touch each other and lend support on a subcontract basis."

In addition to its technology, Bailey said DocuLex brings a strong install base and channel to the table. "From a technology standpoint, it makes sense for Elephant Outlook to point its customers toward our software now for e-mail archiving," he said. "DocuLex also has a lot of customers, which creates a lot of maintenance revenue. In addition, none of the other ISVs has sold through a channel, and we bring a pretty established network of MFP dealers."

Bailey said that as margins shrink on printing sales, MFP dealers are looking to go deeper into managed services like the ones now offered by Protected Trust. "They want to move beyond managed print services, which is becoming a commodity," he said. "We're seeing a lot of dealers looking to acquire IT companies and then scratching their heads on how to make it work. We can help them get there by providing support for network management, back-up, and data recovery services, not to mention document management."

Bailey said that about 20% of DocuLex's customers now have hosted applications. "That area of our business is growing fast, but not that fast," he said. "Inland's data center business may be smaller from a revenue perspective, but from a growth perspective,

it's expanding at a faster rate. This is a very exciting opportunity for us to move into some higher growth markets."

For more information:
http://www.doculex.com/protectedtrust/

Cloud ISV Thrives on Outsourcing Approach

Digitech may be one of the founders of the cloud-ECM market, but don't talk about the cloud as a business driver for the Denver-area-based ISV. No, Digitech views itself as a document outsourcer because that's where the ROI is. According to Sean Morris, director of sales, Digitech's ability to outsource document management applications is what is driving explosive growth in the company's hosted ImageSilo business.

"ImageSilo is adding more and more customers every day," Morris told *DIR*. "The acceptance rate of end users who are willing to store their documents in the cloud is improving. The pendulum has begun to swing our way, and it is still swinging.

"That said, while all the talk in the market is around cloud, we think the acceptance we are seeing in the business community is more around our ability to outsource their document management. If you think about the term 'outsourcing,' it can have a negative connotation. But, at the same time, businesses need to find ways to do things faster, better, easier, and in a more affordable fashion. I think more people than ever are looking to outsource, because it means reduced labor, which has a positive impact on the bottom line.

"So, while the cloud is a hot term, it's the fact that we can show customers use case after use case that has each positively affected a business' bottom line that drives them towards ImageSilo."

AIIM CAPTURE TOOLKIT DUE OUT

Keep your eye out for next week's 2013 Capture Toolkit, published by AIIM. I wrote the introduction, which is basically a summary of why we capture documents. It comes down to two things: data and workflow/archiving. I go over some history as to how these two markets started out separately and then converged—on the high-end at least. Then I also offer some reasons on why this convergence will continue and reach the mid-market and below in upcoming years. There is some other good stuff in there too.

Digitech's deep understanding of document outsourcing goes back to its roots as a developer of image capture software targeting primarily service bureaus. It was through its work with service bureaus that Digitech recently became involved with **OPEX**, a mail sorting, extraction, and scanning hardware vendor out of Moorestown, NJ. OPEX and Digitech recently announced integration between Digitech's software and OPEX's scanning hardware.

"Like us, OPEX really sells the value of its technology and the ROI it can produce," said Morris. "Their scanners are designed to reduce the prep work in service bureaus and that can have a huge impact. Their technology has the ability to reduce labor costs, which is important as the service bureau market continues to become more commoditized.

"When OPEX is out there selling its technology to service bureaus, I want to be in their basket of partners. Their scanners are used to eliminate as much prep work as possible and then scan the images. The images get passed to our software for additional indexing and QC. Of course, our combined technology isn't just for service bureaus, end users can take advantage of it as well."

The OPEX scanners have been integrated with both ImageSilo and Digitech's PaperVision on-premise capture and document management software.

On a technical note, because it was not developed yet when the integration was being completed, OPEX's new "Batch Transform" module, which was released with CertainScan 3.0 [*see DIR 6/7/13*], was not utilized. Rather the Digitech software was integrated with OPEX's previous generation proprietary output format. An updated integration may be done in the future.

For more information:
http://www.documentimagingreport.com/index.php?id=2429
_____

CAPTURE WINS, FROM PAGE 1

be part of the center's opening ceremony."

For more information:
http://www.documentimagingreport.com/index.php?id=2432

A BPO beachhead

ReadSoft's sale of XBOUND was to a BPO provider—which makes sense, as BPOs are where the product has had previous success in Europe and was also tabbed as the first target market for U.S. sales. The buyer is **DISC Corporation**, a Missouri-based organization that focuses on the healthcare market. The deal, which was closed in Q2, is expected to bring ReadSoft $1 million per year.

XBOUND is a platform for managing multiple capture applications and technologies. Its implementation has helped DISC replace manual labor with automated classification and indexing as part of its BetterChart solution for digitizing paper records. DISC is counting on XBOUND's Process Designer to help it rapidly onboard new customers.

For more information:
http://www.documentimagingreport.com/index.php?id=2433