Document Imaging Report

Business Trends on Converting Paper Processes to Electronic Format

4003 Wood Street ● Erie, PA 16509 ● PH (814) 866-2247 ● FX (800) 507-8981 ● http://www.documentimagingreport.com

December 1, 2006

THIS JUST IN!

ABBYY RELEASES FIRST MOBILE SDK APP

Have you seen the commercial in which a guy takes his BlackBerry on his business trip and nothing else? When he goes through airport security, he has no laptop to unpack, no cell phone to take out of his pocket. The point of the spot is that mobile handheld computers are the wave of the future, and from the buzz we've been hearing, that future is arriving rapidly.

With an eye toward this handheld world, at **AIIM 2006**, recognition specialist **ABBYY** introduced its *Mobile OCR SDK*. Last month, ABBYY announced the first application built with this toolkit—*MoBiCR*—a program for capturing business card data with mobile phones. ABBYY built *MoBiCR* for the popular S60 platform, and it has been integrated with the address book software of at least one major vendor. ABBYY expects the application to be available before the end of the year. It will have a list price of $19.99.

"The biggest challenges creating applications for mobile devices is that they have to be small, they have to run within the devices (so users don't waste airtime), and they have to run fairly quickly," said Dean Tang, president and CEO of ABBYY USA. "We have created an OCR application that is only 2 MB in a ROM configuration and 2.5 MB in a RAM deployment. When you reduce the amount of code you use, there is a trade-off in accuracy. We removed technology designed primarily to deal with large amounts of text and maintained the code for recognizing small pieces of text—like you find on business cards.

"In addition, we had to optimize our software to run on the slower CPUs embedded in mobile devices. The CPU in a cell phone cannot even match the slowest Pentium. We know they will be faster in the future, but, we had to make sure our

Continued on page 7

Deadlines Approaching For Claims Forms Upgrades

For the first time in more than 10 years, health insurance companies will be dealing with new billing forms from healthcare providers. After the first of the year, both the HCFA 1500 and the UB-92 are due to be retired in favor of upgraded versions. According to federal mandates, insurers and clearinghouses should already be able to deal with the CMS 1500, with a Feb. 1 cutoff date for using the older HCFA 1500, while use of the UB-04 goes into effect on March 1, with a May 23 deadline for discontinuing use of the UB-92.

Probably the primary difference between the old and new forms is the introduction of a National Provider Identifier (NPI) number that must be included on each form. The NPI is designed to create a uniform code for identifying healthcare providers. Traditionally, different insurers have identified providers in different ways.

"The NPI is like a bar code, in that there is a check-digit feature built into it," said Alan Lattimer, manager of operations for **SunGard** EXP's *FormWorks* data capture software line. "Basically, the check-digit enables insurers to determine, without any database look-ups, if the NPI is a valid number. An NPI can eliminate exceptions created by a misspelled physician's name, for example. All our clients want to take advantage of this." [*Editor's note: The check-digit is based on something called the Luhn Formula, which is pretty cool. Check it out at* http://www.medavanthealth.com/implementation/npi/NPI_check_digit.pdf]

According to Lattimer, the other major change in the CMS 1500 involves splitting the detail lines into two lines. "This changes the location of some of the fields captured through OCR," said Lattimer. "In addition, it is designed to create a dedicated space for supplemental explanations. Historically, this supplemental information could be written anywhere on the form."

This new construct can be a double-edged sword

for capture specialists. "On one hand, it enables us to program our software to automatically locate and capture some of this supplemental information, which can then be leveraged as part of a downstream rules application," Lattimer said. "On the other hand, it also encourages providers to include more supplemental information. Traditionally, maybe 5-10% of HCFA forms had explanations on them, mostly related to something like an anesthesia procedure. With this new format, we expect that percentage to increase, which will increase the complexity of the capture process."

"Basically, [health insurance companies] are so used to a standard set of rules and output, that the new forms are turning some of their processes on their heads. It's not just the capture systems that are affected, but the data output systems as well."

— Alan Lattimer, SunGard EXP

SunGard is one of the leaders in the software market for automated capture of data from healthcare claims forms. Between its *FormWorks* and *Entrendex* product lines, it has more than 150 customers using its software. *Entrendex* is an internally developed solution that leverages recognition technology from **Mitek Systems**. *FormWorks* is billed as SunGard's "next-generation" solution and was acquired with **Recognition Research** last year [*see DIR 8/5/05*].

"At the beginning of 2006, we started looking at how to manage this transition to the new forms," Lattimer told *DIR*. "We started getting preliminary documents and determining what new features we had to add. Basically, to upgrade, we are charging customers for the addition of two new forms to their capture package."

SunGard is also a leader in the imaging, workflow, storage and retrieval software market for claims forms, with its MACESS product line. "In that area our customers are incorporating NPI numbers into their indexing and query processes, as well as their workflow queues," said Stacy Reed, a senior systems analyst for SunGard EXP.

Don't underestimate the effects

It's Lattimer's opinion that many of SunGard's customers underestimate the effect that apparently minor forms changes will have on their business processes. "It's the type of thing where they might call us in expecting the transition to take a few days, and it ends up taking a few weeks, or months," said Lattimer. "Basically, they are so used to a standard set of rules and output, that the new forms are turning some of their processes on their heads. It's not just the capture systems that are affected, but the data output systems as well.

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. 16, No. 23

Editor: Ralph Gammon
4003 Wood Street
Erie, PA 16509
PH (814) 866-2247
FX (800) 507-8981
ralphg@documentimagingreport.com

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

Publisher and Circulation Manager
Larry Roberts
RMG Enterprises, Inc.
5905 Beacon Hill Lane
Erie, PA 16509
PH (412) 480-5116
FX (412) 291-1352
larry@rmgenterprises.com

DIR is published 24x per year, on the 1st & 3rd Fridays of the month, by:
RMG Enterprises, Inc.
5905 Beacon Hill Lane
Erie, PA 16509
PH (412) 480-5116
http://www.documentimagingreport.com

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

"Because it's been 14 years since the last form change, many of the business process analysts at the insurance companies have never dealt with changing forms. One of the common reactions we get is, 'well, I didn't change the form.' Guess what? I didn't either. It's just something we all have to deal with. The key to eliminating some of the frustration for insurers is to plan to look at all their processes, determine how they will be affected, and then go from there."

According to Lattimer, many of SunGard's customers are starting to come around and realize they need to upgrade their claims processing applications. "Part of the problem is that samples of the new forms have been pretty hard to come by," said Lattimer. "However, since the CMS 1500s went into mass production at the beginning of October, we've started to see a ramp-up from our customer base.

"With their current applications, it's still possible to process the new forms, it's just not as efficient. Right now, maybe only 5% of the forms they are receiving are CMS 1500s, and 95% are the old HCFA 1500s. In two to three months, I expect those numbers to invert, and 95% will be the new forms. That's when the pressure will really be on."

Deadline a moving target

Reed added that the Feb. 1 cutoff date for switching to the CMS 1500 is not necessarily set in stone. "That's a federal deadline, but we've seen a lot of insurers go to their state governments and ask for extensions," she said. "As a result, we are starting to see deadlines that are all over the place. The UB-deadline is a little further out, so we haven't seen as much action regarding that yet." [*Editor's note: the CMS/HCFA 1500 forms are submitted by physicians' practices, while the UB-92/-04 forms are submitted by institutions, such as hospitals*.]

Reed concluded that some of SunGard's customers are using the new forms as an impetus to try to increase the number of EDI claims they are receiving. "Some of the new output requirements are designed to encourage adoption of EDI," she said. "We are definitely seeing more customers jumping on the EDI bandwagon. Depending on the state where an insurer is located, there are different levels of EDI encouragement. Currently, the percentage of EDI claims can vary widely from customer to customer, and we expect that trend to continue."

For more information: http://www.macessexp.com/; http://expweb.sungard.com/SungardEXP_Public/default.aspx?id=19; http://www.rrinc.com

Peladon, Wausau Sign OEM Deal

Intelligent document recognition (IDR) software specialist, **Peladon** has signed the first OEM agreement for its recently introduced classification and extraction tools. Transaction processing ISV **Wausau** will introduce Peladon's technology into its *ImageRPS* capture system. It will increase the forms processing capabilities of the historically remittance-centric *ImageRPS* and enable Wausau to more aggressively pursue new capture opportunities.

Like all transaction processing vendors with a paper legacy, Wausau faces the challenge of growing its business despite the increasing adoption of electronic payments. Expansion into related markets like wholesale remittance and EOB (explanation of benefit) processing, as well as ECM, has proven a natural tack for these vendors [*see DIR 2/17/06*]. "As margins continue to decline in traditional remittance processing, we are looking to offer more premium services to drive revenue," said Patrick Brzezinski, remittance product line manager for Wausau. "To maximize the efficiency gains and ROI associated with these services, we needed an automated forms processing package."

Peladon is a San Diego-based ISV, launched in 2005 with a focus on processing semi-structured documents [*see DIR 6/3/05*]. This year, Peladon introduced auto-classification and auto-extraction modules (in part, at least) to fill a hole in the market created by **Captiva's** 2005 acquisition of French IDR specialist **SWT** [*see DIR 7/7/06*]. "We really like Peladon's ability to integrate its technology directly with *ImageRPS*," said Brzezinski. "It enables our customers to leverage IDR without ever having to leave their Wausau workflow or interface. We couldn't do that with any other third-party product."

"Basically our product is comprised of a series of DLL files that can be plugged into somebody else's application," said Phil Bennett, CTO of Peladon. "After working with our software for two days, Wausau was able to launch our processes and reflect what they could do with documents within the *ImageRPS* platform. They were able to integrate both our GUIs and our processes in a way that is fairly seamless to their customers."

Wausau originally introduced a wholesale remittance product in 2004. It leveraged **Nuance** OCR technology and utilized a rope-and-zone process [*see DIR 7/2/04*]. "About 25% of our customers do some sort of wholesale remittance or heavy exception-type processing, taking advantage of that technology," said Brzezinski. "The deal with

Peladon enables us to really start applying automated capture to those processes."

"Our software has the ability to efficiently capture data from the complex tables that you often run into when processing EOBs and wholesale remittance forms [*a wholesale remittance is a corporate payment that typically includes one check to cover multiple purchases*]," said Bennett. "Complex tables are not structured row-by-row, and columns can break depending on what row they're associated with. We've created an easy interface that enables a user to define how a complex table should be processed, and then going forward, whenever that type of table appears again, our software can automatically recognize and process it."

According to Brzezinski, IDR is attractive across Wausau's customer base, which includes 30 Fortune 500 companies and more than 300 financial institutions, including four of North America's top six lockbox operations. "The Peladon technology should help us move more deeply into the wholesale remittance market, as well as into some targeted vertical applications in areas like child support, medical payments (EOBs), and transportation forms.

"The bottom line is that every one of our customers has hard-to-process exceptions or industry specific forms. These documents might be combined with payments, or they might just be captured for workflow and document management purposes. Regardless of the application, we want to help our customers reduce the number of vendors they are dealing with, by offering them a broad-based capture solution. [*Editor's note: Brzezinski noted that 10% of Wausau's revenue is now being generated through its Optima IMS ECM product line—an OEM version of* ***Hyland's*** *OnBase software, and that percentage is growing fairly rapidly*.]

The IDR-enabled version of *ImageRPS* is due to be released for beta testing before the end of the year, with general availability scheduled for March 2007.

Peladon getting established

Peladon is currently in discussions with other potential OEM partners and recently signed on two VARs. The VARs are medical systems specialist **OptiForm** and document imaging-focused **Results Engineering**. "OptiForm has an EOB practice and will use our software to further automate its solution," said Noel Flynn, COO of Peladon. "They already have technology for mapping data to the appropriate back-end systems, which means we can focus on data capture—our specialty. Results Engineering already sells some other data capture products. However, we think we offer a product that is easier to set up and walk away from at end user sites, as well as one with a simpler pricing model than competitive products that carry a click-charge or volume-based pricing."

Peladon also has an OEM agreement with **SunGard** that remains in place. Earlier this year, Peladon was acquired by U.K.-based data capture specialist **DRS** [*see DIR 2/17/06*]. "This year, our focus has been on developing our classification and data extraction products, getting them in use, and establishing reference sites," said Flynn. "We are also continuing to develop our channel. Next year, we will really focus on growing our revenue."

For more information: http://www.peladonsoftware.com/; http://www.wausaufs.com/; http://www.documentimagingreport.com/Peladon_Wassau.1530.0.html

Kofax Introduces IDR For Resellers

With the latest release of *Xtrata Pro*, **Kofax** continues its pursuit to bring forms processing to its wide-ranging reseller base. *Xtrata Pro* is designed to be an easy to implement and use solution for classifying and extracting data from semi- and unstructured forms. It is based primarily on the technology Kofax acquired with **LCI** earlier this year [*see DIR 3/17/06*].

At Kofax's recent European partner event, we saw a demo of *Xtrata Pro* being used as an invoice processing solution. "It's a learn-by-example application that starts with full-page OCR," explained Richard Brierton, product manger for Kofax. "About 20 or 30 samples should be enough to recognize a particular class of document—although that depends on how many classes you want to define and how similar your document types are."

Once a document's classification is determined, point-and-click and drag-and-drop commands can be used to hone in on the data. These commands are memorized, and the next time the application sees that particular class of document, the data capture process is automated. "We are emphasizing ease-of-use and re-use of components," said Brierton. "We are packaging invoice processing rules and commands in the box, but resellers with vertical specialties can easily incorporate their own sets of rules for the distinct document types they deal with. These resellers can then repackage the rule sets in vertically specific versions of the software."

Xtrata Pro is being sold as a module to be used within Kofax's *Ascent Capture* platform. Kofax is positioning it as having capabilities that lie between

the basic data capture included with *Ascent* and the advanced capabilities that come with the company's *Indicius* application. *Xtrata Pro* is being targeted as a replacement for the *Ascent for Payables* product that is based on **Océ ODT's** recognition tool kit, and is often implemented by Kofax resellers to capture header and indexing data from invoices.

"In 2003, we decided that we wanted to own our technology in the data capture arena—what we call transformation," said Sameer Samat, VP of engineering for Kofax parent **Dicom**. "Since then, our acquisitions of Mohomine, Neurascript, and LCI have all increased our capabilities in transformation and document recognition. Our goal is to take this best-of-breed software and leverage it to become a dominant player in this market.

"To do that, we realize we have to be an innovative technology leader. I think we've done a good job in this area. We have taken our technology portfolio and created some very high-end transformation solutions in areas like auto-separation of documents. We've helped organizations accomplish things they never thought they could.

"Our second goal is to simplify the technology, make it as mainstream as possible, and make it available through our channel partners. *Xtrata Pro* is our first effort at this. Basically, we've taken multiple engines that do classification and text extraction and hidden them behind a simple user interface. A few years ago, we introduced *Xtrata* for capturing data from structured forms [*see DIR 2/20/04*]. What we've done now is extend that brand for semi- and unstructured documents."

For more information:
http://www.documentimagingreport.com/Xtrata_Pro.1532.0.html
http://www.kofax.com/products/ascent/xtrata_pro/index.asp

Visioneer Increases Network Options

Visioneer recently introduced a pair of new network scanning options. First, it announced that five of its **Xerox** DocuMate units had been certified to run with the **AXIS** 70U Network Document Server. Second, it announced that its Patriot 470 model had been certified to work with the **Kofax** Document Scan Server (DSS). At **AIIM 2006**, Visioneer introduced its initial network scanning option, a NetScan 2000 box, which can be integrated with all its DocuMate models, as well as its Strobe portable scanner [*see DIR 6/16/06*].

"We are now offering three different levels of network scanning," explained John Dexter, director of business development for Visioneer. "The NetScan is really aimed at giving several people in the same workgroup access to a single scanner. The AXIS is specifically designed for scan-to-distribution applications. The Scan Server is designed for enterprises that want to scan directly into back-end systems."

Since, we've already covered the NetScan 2000 and the Kofax DSS [*see DIR 5/19/06*], we felt the need to delve more deeply into the AXIS 70U, which lists for $499. Like the NetScan and DSS devices, it has a USB plug-in for the scanner and an Ethernet port for connecting to a network, eliminating the need to attach the scanner directly to a PC. The AXIS 70U also features a CPU, which runs a Unix operating system, and an LCD interface for accessing e-mail address books. A keyboard can be attached to the device to increase e-mail options.

"The primary destination for documents captured with scanners attached to the AXIS 70U will be e-mail addresses or watched network folders that can be accessed by enterprise capture applications," said Dexter. "It works great as a replacement for a fax machine and makes scanning much simpler and more accessible than a traditional PC configuration."

The 70U contains the TWAIN drivers for the scanning devices it supports. In addition to the Visioneer models, the unit supports workgroup models from **Fujitsu**, **Canon**, **Epson**, **HP**, **Plustek**, and **Mustek**. It is available through distributors like **Tech Data** and **Ingram Micro**.

"We've found the combination of the DocuMate 632 [departmental model] and the AXIS 70U make a great alternative to HP's Digital Sender," said Dexter. "It may not have all of the advanced features of the Digital Sender in areas like authentication and application integration, but our solution offers superior image quality and better paper handling—especially when it comes to clearing jams. The 632 also has a legal-sized flatbed for exception documents. Plus, the street price of a new Digital Sender is around $2,500, while the 632 plus the AXIS 70U sells for around $1,700."

Swedish healthcare provider **Diversified Specialty Institutes** is in the process of rolling out a scanning system based on the AXIS 70U and Visioneer scanners. "They were upgrading from a network of older Digital Senders and chose our solution rather than the latest Digital Sender model," said Dexter. "They are using our system to scan medical files and billing information."

DocuMate DSS approval on way

Visioneer is currently waiting to receive approval

from **Xerox** on the integration of its DocuMate devices with the Kofax DSS. According to Murray Dennis, president and CEO of Visioneer, Xerox needs to grant approval before Visioneer can market any network solutions using its brand name. Visioneer signed an agreement to market its scanners under the Xerox DocuMate brand in 2003 [*see DIR 9/5/05*].

"We had to get similar approval for the NetScan and AXIS devices," said Dennis. "However, integration with those devices was a little more straightforward than DSS integration. It did not involve writing new drivers [*DSS bypasses traditional drivers by enabling applications to connect with scanners through Web services calls*]. We're talking about a difference of months vs. weeks. We expect to gain DSS approval across the DocuMate line sometime in the first quarter."

The Visioneer-branded, DSS-approved Patriot 470, is essentially the same machine as the Xerox DocuMate 262, which indicates that the technical work for integration with the DocuMates is already completed. The Patriot line was launched by Visioneer earlier this year, specifically to meet the manufacturing requirements of GSA listings with the federal government.

For more information:
http://www.documentimagingreport.com/Visioneer_Axiom.1531.0.html
http://www.axis.com/products/axis_70u/index.htm;
http://www.documentimagingreport.com/Visioneer_DSS.1536.0.html

Iron Mountain Digital Shows Strong Growth

Iron Mountain continues to grow its digital business—expecting to generate some $135 million in digitally-related revenue in 2006. This may seem like a drop in the bucket for a $2 billion company, but it represents significant success for a practice that was launched just five years ago. And in accordance with Iron Mountain's corporate goals, the Digital Business Unit has succeeded in establishing itself as a top-two player in both its areas of focus. It is also generating a sound base of recurring revenue.

"The goal of our Digital Business Unit has been to mirror our success in the physical storage world," said John Clancy, Iron Mountain's executive VP of digital business. "In the physical realm, we focus on storage of paper records and back-up tapes. In the digital realm, we've chosen two rails that run along the same lines: electronic records and online data back-up and restoration. Focusing on those two areas, we've managed to surpass the $100 million annual revenue mark, reach profitability, and achieve a base of 80% recurring revenue."

Iron Mountain first launched its digital business initiative in 2001 [*see DIR 9/6/02*]. "At that time, we had a broadly defined offering that included capabilities like COLD and dedicated document imaging," said Clancy. "But, we found many records in those areas are much more active than the types of records Iron Mountain typically deals with. We're used to keeping paper in a box and delivering it to a customer when they have a need for it. Most document imaging systems have a more specialized workflow."

E-mail archiving became a natural place for Iron Mountain's digital focus, especially in the wake of the 1998 implementation of **SEC** rules 17a 3-4 and a subsequent series of million dollar-plus fines levied in 2002 for violations of these rules [*see DIR 12/20/02*]. "Our primary focus has been on the broker/dealer community, whose e-mail archives are regulated through SEC 17a 3-4," said Clancy. "We have more than five billion e-mail assets in storage, and our customers include some of the largest firms in the financial services industry."

Now, with impending changes to the federal Rules of Civil Procedure, which will regulate the discovery of electronic records for court cases, concerns about e-mail and other types of electronic records archiving are spreading to markets outside the broker/dealer community [*see DIR 1/20/06*]. "To date, e-mail has been the primary electronic record of significance and is where the budget dollars have been spent," said Clancy. "Focusing on e-mail has helped us establish our beachhead in electronic records management."

To help it expand into other types of e-records, at the recent ARMA show, Iron Mountain announced a partnership with records management specialist **Accutrac Software** and previewed a new retention management solution. "It's still early in the e-records management game, and there is a lot of opportunity," said Clancy. "Over time, we think it will be a bigger space than online back-up and restoration. Of course, unlike back-up and restoration, which is pretty horizontal, there are a lot of vertical market issues to consider when addressing e-records. This is something we are looking into."

Distributed back-up leader

While Iron Mountain's e-mail archiving business has grown organically, utilizing a solution based on **EMC's** *EmailXtender* software, its online back-up and restoration business has been driven by a pair of

acquisitions. In 2004, Iron Mountain acquired Connected, which specialized in backing up and restoring distributed PCs. In 2005, it acquired LiveVault, which focused on server back-up. "That combination has vaulted us to a leadership position in the market for distributed data protection as a hosted service," noted Clancy, who came to Iron Mountain with the Connected acquisition. "This includes data that sits on the edge of a network—on PCs and/or remote servers. A lot of CIOs spend their time thinking about mission critical applications, but the reality is, the majority of an enterprise's data sits outside its tier-one data center. According to **IDC,** in fact, 60% of enterprise data is 'distributed.'"

Back-up vs. archiving

Because it offers both services, Iron Mountain has a unique perspective when it comes to the ongoing debate between back-up and archiving. "If you were to poll 100 CIOs and ask them to differentiate between the two, I think you'd get 20-30 different answers," said Clancy. "We have some customers that use a mathematical formula to compare the cost of a true archiving solution vs. the cost of restoring files from back-up tapes. They factor in a percentage that reflects the risk of litigation that would require pulling old files."

Unlimited growth potential

Iron Mountain's 2006 numbers represent 50% growth in the company's Digital Business Unit over 2005, which was another year of 50% growth. Of its customer base of 200,000 businesses, less than a quarter currently use Iron Mountain's digital services. This includes 4,000 corporate and 40,000 SMB customers. "Even within our current digital customer base, the growth potential is enormous," said Clancy. "The market for e-records management and data protection can be measured in the billions. Our future direction will be determined by which digital records are most valuable to our customer base."

We concluded by asking Clancy if Iron Mountain had any more acquisition plans to help it maintain to the current 50% growth rate of its Digital Business Unit. After all, the company has been famously acquisitive in its physical storage business. "Getting software code from several acquisitions to work together is a lot more challenging than combining real estate and customer lists, as you do in the world of physical storage," he replied. "So, no, we will not be as acquisitive on the digital side. That said, like most technology companies, acquisition is part of our strategy.

"LiveVault and Connected were both partners before they were acquired, so that's proven a route for us to take. The bottom line is that, if an acquisition can help us meet a strategic goal and accelerate our time to market with a particular product or service, we will consider it."

For more information:
http://www.ironmountain.com/digital/

ABBYY MOBILE, FROM PAGE 1

application was fast and accurate enough to work now.

"Finally, there are image processing challenges to take into account because of the unique lighting and curvature effects when working with cameras on mobile devices. We are starting to see some start-ups that focus specifically on processing images from mobile devices."

One of those start-ups is Palo Alto, CA-based **scanR**, which licenses ABBYY's traditional OCR technology. scanR makes applications for business card, white board, and document capture available to its customers as a service. scanR's management team features veterans from **HP**, **Kodak**, **IPIX** and **Intel** [*see DIR 4/7/06*].

Recognizing the future

"I just got back from a wireless conference in Monterey, where I saw fascinating new technology surrounding digital camera devices," said Tang. "For example, I saw lenses that enable multiple focus points, so not only does the object in the center of a photo show up clearly, but so do the foreground and background objects. I expect that by 2010, this type of lens will be used on all camera phones.

"One interesting trend was that a lot of the technology was developed by Silicon Valley firms, whereas traditionally, digital camera technology has come primarily out of Japan and Germany. I think that says something interesting about the future direction of the mobile camera space."

The next level of distributed capture

MoBiCR, captures information from business cards in less than five seconds. "The quickest I've seen is 1-1.5 seconds, and the slowest is 5-6 seconds," said ABBYY's Ilya Evdokimov, who gave us a demo. "Typically, we ask that at least a 1.3 megapixel camera be used. At that level, you can expect around 99% field-level accuracy. With 1 megapixel, your errors start to increase.

"Of course, you have to consider lighting, from how far away a picture was taken, and focus. The auto-focus on most smartphones takes care of the

last issue. Basically, if you can read the words on the image with your eyes, our software should be able to read them. Our software also makes a decision on whether to use a color image or convert it to black-and-white before applying recognition. Color adds depth and can improve accuracy."

"If you look at the flow of capture from the back-office to the distributed desktop, the next natural step is very possibly mobile devices."

— Dean Tang, ABBYY USA

The demonstration we saw gave the user a chance to verify captured data, make any corrections, and release it to an address book. ABBYY's *Mobile OCR SDK* will work with almost any operating system. However, mapping captured fields to address books requires cooperation with phone vendors.

"We are showing our SDK to all the cell phone manufacturers," said Tang. "Whenever we release a new technology, we always pick a showcase application that we develop ourselves. This gives our potential customers an opportunity to see the power of the technology in a dedicated environment. It would not surprise me if, within two years, every cell phone targeted at mobile professionals includes some type of OCR."

Tang indicated that both start-ups and traditional capture vendors are looking to leverage cameras on mobile devices. "One start-up licensing our technology is **ActiveSymbols** [http://www.activesymbols.com]," said Tang. "They enable you to do things like take a picture of a CD cover in a store, apply OCR to determine the artist and title, and then run a price comparison against online stores like **Amazon**."

Tang concluded that he views capture from mobile devices as the next natural step in the evolution of document scanning. "Ten years ago, the market was dominated by centralized, high-speed scanners," he said. "In the past five years, the high-growth area has been workgroup devices. If you look at the flow of capture from the back-office to the distributed desktop, the next natural step is very possibly mobile devices.

"Taking it a step further, I recently saw HP demoing a wearable camera. It was designed to be active 24x7 and worn like eyeglasses. The idea is that all this video would be uploaded to a server farm. To find what you wanted, the video would have to be made searchable, which seems like a great opportunity for us. In the more direct future, by 2010, we expect there to be one billion camera phones in circulation (there are now some 200 million). If you compare that to a market of less than 1 million document scanners, you can see some of the potential for the *Mobile OCR SDK*.

For more information:

http://www.documentimagingreport.com/ABBYY_Phone_OCR.1527.0.html
http://www.abbyy.com/press/press_releases.asp?param=56761