HARVEY SPENCER ASSOCIATES

Intelligent Capture

How this Benefits YOU

Sponsored by: opentext™

Introduction

Intelligent Capture consists of a series of technologies including pattern recognition such as Optical Character Recognition (OCR) supplemented by AI techniques, computer vision and deep learning with rules based validations to understand, classify, extract and label information from a broad set of incoming multi-channel input information. Inputs that help control or support the business process or compliance needs come from many sources in a multitude of formats. Business leaders must manage and take control over relevant sources of information and most of this is unstructured.

Data can come from documents, mobile devices, voice conversations, chat sessions social media, or images that may be distributed and utilized throughout the organization. Intelligent Capture is a software set of solutions that allows these multi-source inputs to be understood, classified, validated and where appropriate, provide relevant extracted data for automatic entry into the business process.

Intelligent Capture Drives Business Value

A large German based multinational company started a "paperless office" initiative with a single capture application for H-R to just create electronic images of all personnel files, saving floor space and integrating with their ECM and SAP systems. Through the success of this project and the demonstrated benefits, capture and automation gained management's attention. The company continued to realize these benefits as they added multiple business application without adding additional personnel. By utilizing Intelligent Capture, they achieved greater accuracy, and had a solution that was easier to deploy and implement in real-time to drive value. In their shared services environment they found it advantageous to capture all incoming documents for all business units.

Intelligent Capture Services reduce cost and enable new business processes and capability to be invoked through enhanced classification integrated with the business needs and extraction of and tagging or labeling relevant validated data.

Why Hasn't Every Organization Done This?

There are many companies who have implemented basic capture to convert and index their paper to solve a departmental problem and stopped there. Often it was in response to an immediate need:

- "We have purchased an ECM solution to streamline our workflow, we need to install a basic scanning system to get the paper into a digital format"
- "We need to reduce our paper storage costs, so we installed a batch scanning solution to convert our back files of paper into indexed images"

Having successfully installed and used a scanning solution for a number of years, organizations have achieved the first step in realizing the "paperless office", but often do not think about how to improve or expand their solution. The initial solution is out of mind from the IT department and CIO, who have achieved initial ROI payback goals and have other pressing issues to worry about. Here is an outline of relative costs of a end-to-end scanning operations with limited automation:

Capture Step	Labor Cost	Comment
Document Preparation	+++	Clerical
Scanning	+++	Scanner Operator Cost
Recognition	0	Unattended
Indexing & Validation	+++++	Very Labor Intensive
QC and Exception Processing	+++	QC and Rescan Cost
Release to Workflow	0	Unattended

Document preparation, indexing/ validation, quality control and exception processing are all labor intensive. These costs can be reduced through application of Advance Recognition technology. In addition to cost savings the speed of processing can add substantial value to Line of Business applications and shared services environments. Now is the time to look into how Intelligent Capture technologies can help improve processes and reduce costs.

Value of Intelligent Capture

Capture, Manage and Control Multiple Inputs

Capture is no longer just about processing scanned images. In the state of the art digital workplace information is begin captured, analyzed, synthesized and utilized

to create and enhance business value. Today's requirements include the ability to process images created via faxes, mobile devices, digital copiers (MFPs) and electronically created images. Once data from incoming images is recognized, extracted and routed, these data can be called for utilization in business processes and analytics. Given today's digital workplace why hasn't every organization embraced Intelligent Capture to automate business processes?

Most organizations are already running a batch oriented capture system, which has resulted in efficiencies in terms of lower cost, faster, more efficient capture and better management of paper. Surprisingly many capture solutions are still being bought and used for the limited purpose of scanning and manually indexing paper and electronic documents. In 2019 we estimate that $300m is still being spent on these types of basic capture solutions. As these basic solutions often run in the mailroom or back office they often receive little "C Level" attention. The prevailing thought is often '*oh well it will all be electronic soon since paper is going away'* and as the saying goes '*if it ain't broke, don't fix it*'. While expecting big growth in mobile capture, electronic mail and social media input, the reality is that paper is still a very important part of business processes. Organizations are being called on to generate actionable data to complete transactions and for compliance, analytics and customer centric activities -- without inputting data from paper records, it is incomplete.

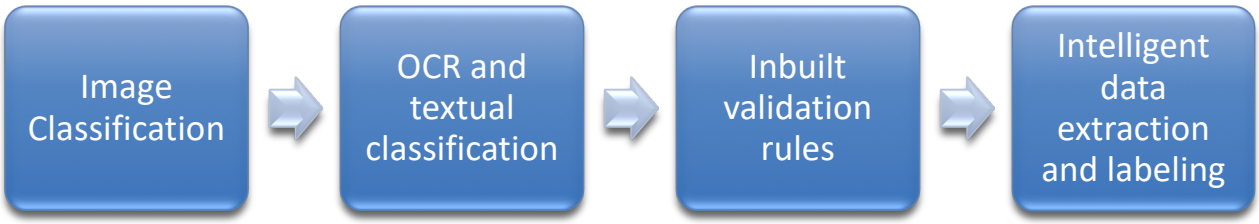
Although scan and store is great and you're saving money, the fact is that you are only scratching the surface in terms of how capture technology can help your organization. Digital Transformation is incrementally increasing the amount of image and unstructured data that affects your business. Daily basic scan-and-store will not help you manage this new influx of data. The only way to adapt to new market digital workplace requirements is with Advanced Recognition. In 2018 over 53% of Capture Software spend was on software to capture paper using advanced recognition to extract meaning.

Automated Classification Drives Business Value

In the past, semi and unstructured documents had to be sorted with separators entered between document types. Manual processes were slow as operators had to insert patch cards or separator sheets to classify documents, files and folders. These labor-intensive processes add cost, require printing, and time to insert and potentially remove them after scanning.

Consider how human understanding is applied to processing a document; the operator picks up a document, and quickly identifies the document type. Once a knowledge worker recognizes the document type he or she can then bring up an interface to manually extract data. The system validates the data against rules associated with the field using a database lookup. If you get it wrong, then it must be retyped. It's all very time consuming.

Intelligent Capture works in a similar way but more efficiently. The automated capture system looks quickly at the size of the image, overall layout and then using this information, determines the document type. This level of automation provides the high level of routing and indexing information necessary to decide what comes next.



Capture solutions with advanced classification leverage both image and textual elements to understand the document type then locate and identify the needed data to extract. In the case of printed paper documents, this saves the printing costs of barcodes, patch codes, and separator sheets and reduces or even eliminates the need for manual sorting. With today's improved feeders in document scanners, the cost of document prep, which was previously around $10 per thousand pages, is reduced to nearly nothing. Most organizations report a return on investment within 12 months simply by implementing automated classification alone.

More business applications are now able to take advantage of data that is generated in field offices, customer facing encounters and even information coming direct from customers themselves. On Demand Transaction Capture has been growing 10-15% per year and we see this growth continuing in the near future. Mobile capture is growing at around 20% annually. Integrating proven Intelligent Capture capabilities within a remote/distributed capture workflow can be effective for some digital workplace applications.

Document Classification Reduces the Cost of Set-up

In basic transaction capture solutions, managers spend large amounts of time setting up templates of new form types so that operators can manually scan a batch of the same document type. In order to identify fields the operator has to label or barcode relevant areas of the form. They then have to laboriously go field-by-field drawing around each zone with a mouse and then connect that with validation rules. Some simple rules are sometimes included, but often the look-ups and validation rules have to be specifically programmed.

Intelligent Capture uses Artificial Intelligence (AI) techniques to "learn" new document sets and other input types, which then can be incorporated into the rule set. Intelligent Capture technology works on the basis of auto analysis of textual document content, image-based analysis of document topology or a combination of textual and image analysis. The system can be trained to identify different document types and subsequently "learn" to identify variations. This reduces the need to manually set up and identify each specific document type and locate the fields to capture. Small changes between similar documents automatically become incorporated, eliminating the need for customization and coding.

A digital mailroom is a good example of how Intelligent Capture is being applied. Mail is opened, prepped and scanned. Electronic documents can also be input via email or watched folder. The Intelligent Capture classification then goes to work to identify the document type. There of course can be exceptions flagged that require verification by the operator and could provide the basis for further recognition training. OCR, ICR (Intelligent Character Recognition), barcodes and OMR (Optical Mark Recognition) are used for auto indexing, extraction or routing. Uncertain fields are verified either automatically or through some human intervention, then the documents and associated data are sent to relevant databases. Documents can then be called by workflows for routing to appropriate recipients for review and approval.

Speed - Extract and Validate Data Faster

Digital Transformation means that business velocity ever increases -- documents have to keep up. Through Intelligent Capture, the system can "classify a document" but it can also 'read a document' quickly. OCR is not new technology, but it has been enhanced and optimized for various business applications.

Application specific lexicons, business rules and the ability to interpret rich image data have improved accuracy rates that can be employed as needed for particular business applications. Application oriented OCR is being applied to invoices, forms, shipping documents, applications and other areas for the automation of business processes.

Validation of the extracted data is important in making the system reliable. Using tabular forms and invoice processing as examples, solutions utilize mathematical formulas such as quantity x unit price to validate the extended amount, as well as calculate discounts. If barcodes provide information concerning document type, then the format must be identified and the barcode converted to relevant information. If the form is a survey, using OMR (Optical Mark Recognition) or check boxes recognition can be easily automated. By leveraging an understanding of preset form types (for example invoices that have already been setup) the system can integrate a new document or form type in a fraction of the time that legacy capture and OCR applications required. With Advanced Recognition technology it is now easy to add new document types to the process whether they are on paper, fax, image, or PDF formats. Intelligent Capture enhances auto-validation protocols, improves data accuracy and reduces the number of costly exceptions.

Initiate and Validate Workflows

Utilizing rules and auto-understanding of the incoming information, relevant metadata can be created and added to the images, which in turn can affect workflow. The classification and the extracted validated data, combined with business rules, can result in the generation of relevant metadata that is not on the form while capturing it. One example is an order that exceeds the supplier's credit. It is possible via look-up into the supplier database to place the customer's credit limit within the metadata, and even, provide a look up linkage to his account and history. By capturing all relevant business data immediately, organizations interact with clients and prospects, engaging in real-time.

Within the ECM, ERP or other business application, business processes can be accessed and modified at the time of capture in real-time in order to better understand and route information and images. In the above situation, breaching a credit limit may require a supervisor's approval. So the workflow is modified to allow routing to the supervisor with an attached message, which is automatically created via the capture solution.

Consider a Customer Experience Management (CX) environment where the customer is required to submit a document maybe while on a chat session or phone call to a customer service representative. Using advanced capture the customer may take out a mobile device or scanner to convert a piece of paper to an image on demand. This image could then be displayed the representative's screen, as well as, the customer's with relevant extracted data with perhaps an dynamically updated application or claim form. If the information runs up against a business rule requiring an exception, the workflow routes the new information which may then require an additional approval level.

Take Action

We have discussed the advantages of moving from simple capture to Intelligent Capture systems that are impacting business. These benefits include:

1. Streamlining existing scanning operations – utilization of Intelligent Capture to automatically identify different document types, eliminating the need for separator sheets with the ability to auto-classify documents. This reduces scanning set up cost, and improves efficiency.
2. Understanding and Extraction of more relevant data. With Capture we can "machine read" documents and utilize OC and XML tagging to extract important data. This allows for data validation, improved accuracy and a more effective process.
3. Extraction of data from additional image sources such as scans from mobile devices, faxes, pdf's, email attachments, and other documents as the business application requires. Now, the information from more than just scanned documents is available for input into business processes. PDF documents can be very versatile "containers" of information. By applying Intelligent Capture tools with understanding and metadata tagging, more relevant, accurate information can be extracted, transmitted and utilized.
4. The additional understanding that is obtained through Intelligent Capture opens up opportunities for better understanding and data availability. But through application of AI, it allows for continuous learning which builds understanding and needs from new forms types, creating new data while providing the ability to initiate business workflows when appropriate. Data derived from Intelligent Capture can be utilized in rules based systems that call for managerial intervention when necessary, make data available on demand for customer facing applications, or automatically routes information to multiple business processes.

Intelligent Capture unleashes information that is locked into unstructured and semi-structured inputs opening up opportunities for organizations to improve operationally; provide enhanced customer experience; and differentiate in the marketplace. It is said that "information is power" --- now we have Intelligent Capture to harness that power.