HARVEY SPENCER ASSOCIATES

Transform Multichannel Data into Information with Capture

Sponsored by Hyland Software

Hyland®

NAVIANT®
Process focus. Excellence.

Executive Summary

Documents and semi-structured information are crucial to digital business; they serve as input for processes and are a vital source of insight. Many times, however, the data remains locked in paper documents or in "dumb" static images. It is often seen simply as historical data for records keeping or other archival purposes. Consider that some of this data, such as patient records or insurance claims, must be unharnessed to provide the complete picture and full understanding for decision making. Increasingly, content is created, accessed and submitted into a process from mobile devices rather than from PCs or scanners. In response to changing business dynamics, capture software has evolved to enable data from any input channel to be transformed into actionable information. With Capture 2.0 software and services, multi-source data (from paper to rich media) can be classified, validated, understood and, where appropriate, extracted for automatic entry into a business process or application.

The Capture Software market has grown at double-digit rates since 2016, due to an improved worldwide economy, greater demand for capture technology to support digital transformation and to feed analytic systems, and growth in Cloud Capture Services. Robotic Process Automation (RPA) is also having a positive impact on the demand for Capture Software in answering the need for cognitive automation. As Capture Software becomes more deeply integrated into business processes, however, it is increasingly being optimized and packaged to address the needs of specific vertical industries as well as horizontal applications.

Introduction

In a data-driven economy, companies face increased demand for real-time insights into all incoming data. Many times, however, that data remains locked in unstructured and semi-structured content and must be unharnessed to provide the complete picture and full insights for decision making. The ability to have automated capture, content understanding, identification, and relevant data extraction for use in today's business systems and applications is critical.

> "*Business is getting transacted at ever-faster speeds and automation with AI and robotics is critical to keeping up.*

Capture Services help with the classification and extraction of accurate relevant transactional data, which is then integrated with the business process." - HSA, WW Capture Software Market Report

From its humble origins as a means of converting paper to searchable scanned images, capture software has evolved into a complex set of software tools that are applied to understand, extract and validate data from a variety of information types. Capture is no longer solely about processing scanned images of paper documents; capture technologies are increasingly being applied to content that originated electronically. Companies are not simply looking to extract data from fixed, known form types, rather they are dealing with extracting information from variable forms to support business transactions. The new wave of capture technology will enable organizations to classify, interpret and understand incoming data from multiple channels and transform that data into information that feeds business processes and supports decision making. While back-office processes will still drive demand for capture software, it will also increasingly be used in customer-facing applications where the real time understanding of various forms of digital content is critical.

Gaining an understanding of how the Capture Software market is evolving, especially as it integrates with adjacent technologies, is imperative. Organizations must move beyond just looking at capture as a way of eliminating paper and reducing staff and labor costs. It is a key technology enabler for companies and government agencies as they respond to increasing global regulations and privacy rules, embark on digital transformation initiatives, engage customers and employees, or seek competitive advantage.

Current State of the Capture Software Market

After years of flat growth, the capture market is experiencing a resurgence, driven by revenue associated with cloud-based systems, integration with transactional systems; and cognition services. The worldwide capture software market hit US$3.6B and grew at a rate of 10.4% in 2016, marking the first time in five years that the growth rate reached double digits. HSA forecasts that the market will exceed the US$4B mark in 2018 and US$5.7B by 2021 (see Figure 1. Capture Software Regional Growth, Worldwide).

Figure 1. Capture Software Regional Growth, Worldwide
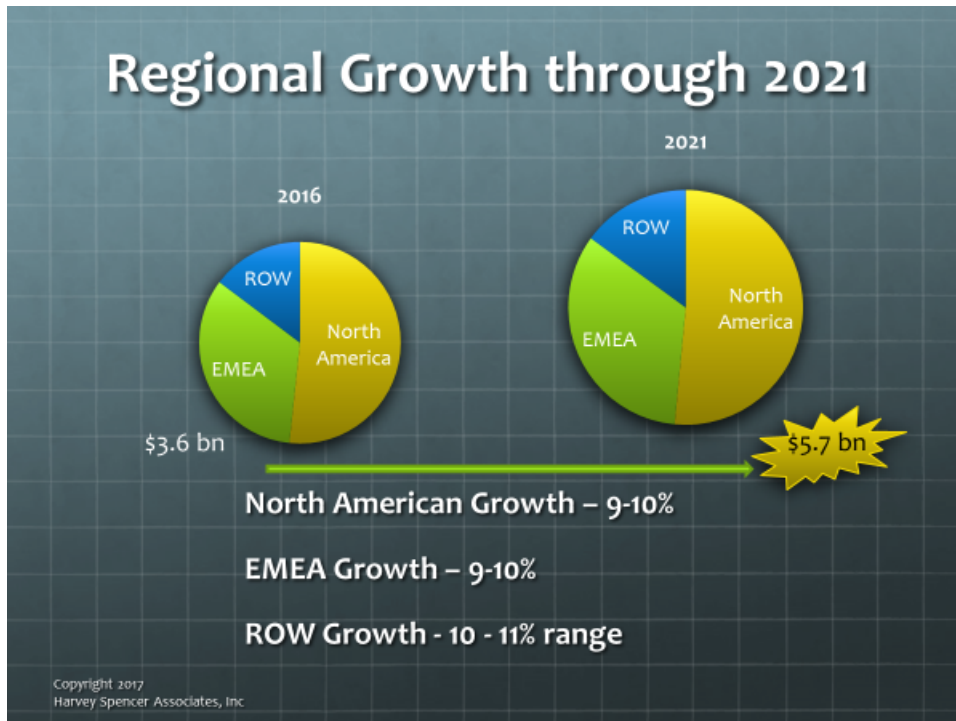
Source: Harvey Spencer Associates. The 2016-2017 Worldwide Capture Software Market Report

What has driven this renewed growth? An improved worldwide economy, greater demand for capture technology to support digital transformation and analytic systems, and the initial growth of Cloud Capture Services drove the double-digit growth. Though Cloud Capture Services still represent a small percentage of the overall Capture market, HSA predicts the growth will be +45% per year through 2021. Cloud enables new means of deploying capture technologies, which will create new opportunities for affecting business processes.

From a technology perspective, Capture Software is also experiencing an evolution driven by improvements in recognition and extraction capabilities as well as intersection and integration with emerging related technologies. In conjunction with business rules, Capture software makes sense of and extracts relevant validated data from a variety of unstructured inputs. Technologies that play a role in the evolution of capture include Optical Character Recognition (OCR), Natural Language Processing (NLP), semantic understanding, and sentiment analysis.

Scanning and basic capture technologies are mature and have been adopted in many government agencies and companies in a variety of industries. Organizations, however, often capture paper documents simply as "dumb" images, with little (or only poor) metadata. While this might provide value for archival purposes or records management, it does little to reduce labor costs and automate processes. In addition, document preparation, indexing/ validation, quality control and exception processing are all labor intensive.

RPA is being implemented to reduce manual keystrokes for simple data transfer. Capture technologies assist by providing document classification, data identification and extraction taking RPA to the next level of automation for some processes. Organizations look to automate their document-centric processes and redeploy staff to activities requiring decision making; capture products have evolved in response, offering better capabilities for classification, data extraction and validation.

Recognition technologies, typically pattern based (e.g., optical character, intelligent character, mark sense and bar code recognition), can be used to make scanned images machine-readable and searchable. The accuracy and speed of these technologies have improved dramatically over the years. Pattern recognition technologies, however, never provide a 100% accurate reading of all the information contained on the document or form. Template-driven OCR was the norm but was prone to high error rates and required text to be in defined locations.

Many capture software products have begun and will continue to use "Artificial Intelligence (AI)" techniques to improve the accuracy and depth of scanning results. The more robust capture engines use neural network technology, a type of machine-learning algorithm, to automatically extract data and learn where to find it, based on document type. Some also use other machine-learning algorithms to assist with the understanding, classification and validation of unstructured data.

What's Next for Capture Software?

The capture software market is undergoing an evolution in response to changing business requirements, organizations' digital transformation initiatives and emerging technologies. HSA refers to this next generation as Capture 2.0 (see Figure 2. Capture 2.0: The Next Generation). Capture 2.0

systems consist of a series of Cloud-compatible RESTful services that use AI and advanced classification to understand which processes any incoming data involves. Such a system can identify duplicates and drive workflows and business intelligence by extracting and/or generating relevant metadata and transactional data from multichannel inputs.

Core elements of Capture 2.0 are as follows:
1. Services Based using a cloud architecture whether on-premise or as a shared service
2. Accepts inputs from multiple sources and is delivered via multiple channels
3. Integrates with Business Processes and can provide info for Analytics and Compliance
4. Services will be called by applications, may be from multiple vendors and response needs will vary

Figure 2. Capture 2.0: The Next Generation
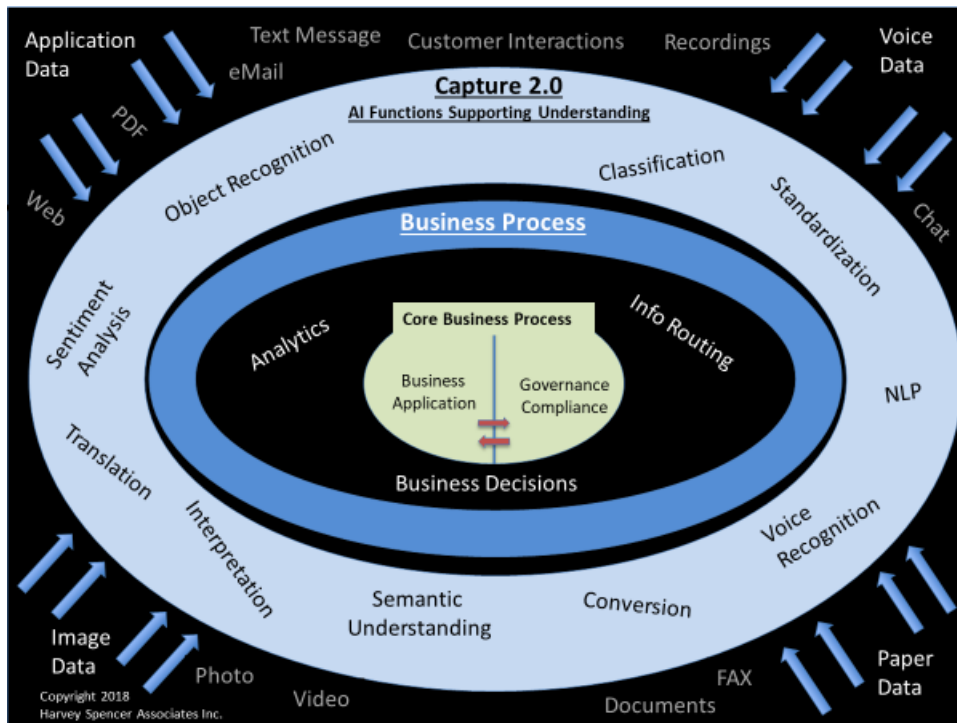
Source: Harvey Spencer Associates

Business applications today can take advantage of data that is generated in field offices or through customer facing encounters. Information that must

© 2018 – Harvey Spencer Associates Inc.

be processed is also coming direct from the customers or constituents. Organizations will still need to process documents that are converted to electronic, such as forms, PDF files, e-mail attachments and scanned documents. Increasingly, however, incoming data may not be textual; it will be still images or photos, video and audio clips and perhaps other forms of data. Input media can be in the form of images (document, photographic) voice, video, text messages (SMS/Chat/Social Media). Mobile devices provide some capture capabilities via their built-in cameras and microphones capable of generating unstructured data combined with relevant metadata. HSA expects continued growth in mobile capture, electronic mail and social media input to business processes. This requires that capture technologies and techniques evolve to handle a diversity of data types and input channels.

Capture 2.0 covers basic imaging to the more advanced sentiment analysis and semantic understanding. Specific technologies that comprise Capture 2.0 systems include, but are not limited to, OCR/ICR, barcode recognition from image, optical mark recognition (OMR), object recognition, voice recognition, natural language processing (NLP), semantic understanding, and sentiment analysis. NLP enables the processing of natural language data using machine-learning algorithms. Semantic understanding, which enables capture software to understand the meaning of textual information, improves classification. Capture 2.0 software engines not only can read and classify a document or piece of unstructured content, but when augmented with sentiment analysis, they can identify and categorize opinions expressed. Sentiment analysis typically uses NLP, text analysis, computation linguistics and biometrics. Further out on the horizon, we expect to see technologies such as video classification, image classification, and voice to text becoming extended services that augment capture and business processes.

Multichannel input and intelligent capture technologies are important, but equally important is their delivery as a RESTful service. Leveraging cloud services as a platform enables integrators to use Capture 2.0 technology to rapidly develop custom business process solutions. As a result, midsize organizations to take advantage of applications that were formerly only available to large enterprises.

Capture 2.0 is as much about business processes and business rules as it is about technology. The technology components are designed to understand the process inputs and through execution of business rules, deliver appropriate data to these processes with limited or no human intervention. Business Rules Processing determines what actions are needed, drives dynamic workflow and decides how best to route information to the proper business application.

Capture Drives Business Value

Capture technologies have often been underutilized, with many organizations simply capturing images of paper documents with some metadata for archival purposes. Beyond scan and store, capture software provides real business value. It helps organizations reduce the labor costs, improve time to cash, support compliance and disaster recovery initiatives, and reduce the cycle times associate with transactional processes. HSA sees an ongoing shift away from records management and archiving use cases for capture towards more transactional and customer facing use cases.

Increasingly, organizations are experiencing demand for real-time insights into all incoming data, as well as the need to unharness and understand the data embedded in unstructured and semi-structured content to support both internal and constituent/customer-facing processes. Capture 2.0 software with voice recognition will foster understanding of the importance of the caller, the nature of the query and what resources may need to be put in place to respond. Using language translation to augment basic capture, will enable companies to process foreign invoices or orders automatically. Applying semantic understanding techniques to incoming correspondence can help improve customer service outcomes. Organizations can better manage customer relationships by extending capture with sentiment analysis.
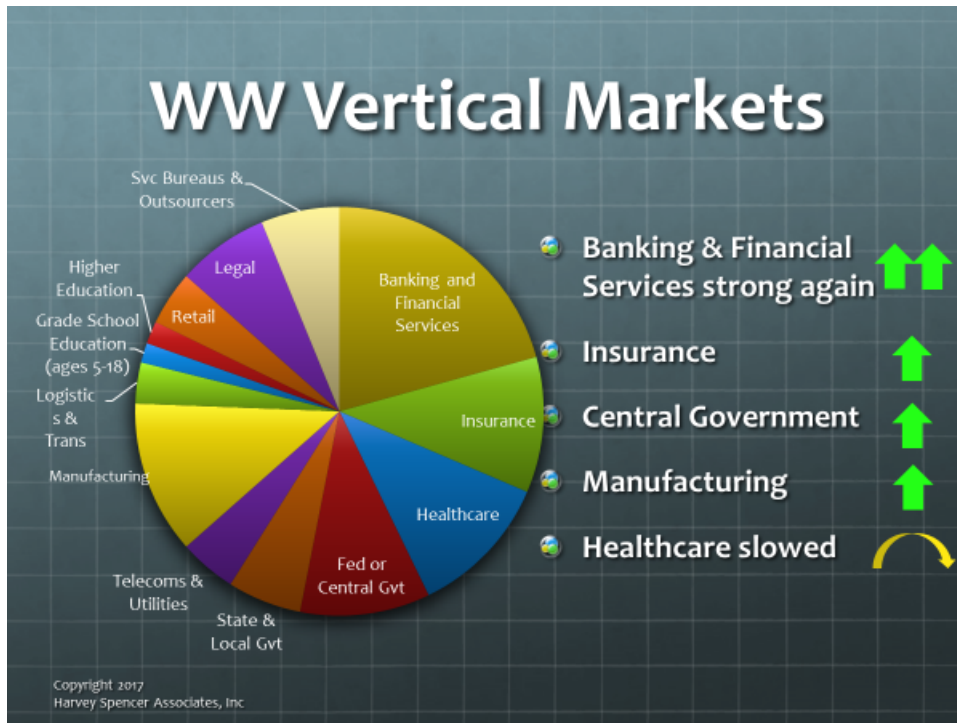
Focused Solutions

Just about every type of industry has a need for some form of capture capability given how much data is still locked in text-based documents, whether electronic or in paper form. Capture 2.0 technologies are designed to classify and extract relevant information for use by structured business processes. As capture software becomes more deeply integrated into

business processes, however, it is increasingly being optimized and packaged to address the needs of specific vertical industries as well as horizontal applications. Packaged solutions address cross-industry applications like financial processes such as accounts payable and accounts receivable; HR processes such as employee onboarding/offboarding; and Supply Chain scenarios for processing quotations, orders, shipping documents and invoices.

The dominant vertical markets leveraging capture software now and in the future are Banking and Financial Services, Insurance, Government and Manufacturing. Healthcare, while still a mainstay, has seen some slowdown in growth (see Figure 3. Worldwide Vertical Markets for Capture). Capture software has also been used in applications within several other markets, including Telecoms and Utilities, Logistics and Transportation, Higher Education, Legal and Retail. Healthcare, which HSA defines as patient facing as opposed to back office claims processing (healthcare services providers, rather than insurers), grew only at 4 percent in 2016. This is in part due to decreased demand in traditional paper-based capture for back-office processes. Though small healthcare practices may still drive demand here. Yet, healthcare administrators find that the need to make healthcare decisions based on data is increasing. Backfile conversion for analytics purposes represents an opportunity. The classification of diagnostic images using neural networks as a means of assisting physicians is another potential application area. In the future we will likely see industry solutions built with capture technologies and extended or related capabilities, such as video and voice classification and analysis.

Figure 3. Worldwide Vertical Markets for Capture

Source: Harvey Spencer Associates. The 2016-2017 Worldwide Capture Software Market Report

Some capture software vendors have built up expertise in these markets and are beginning to offer packaged solutions to address them. Specific solutions include, but are not limited to, those targeted at mortgage lending, medical records, claims and insurance billing. HSA believes these vertically oriented solutions will be enabled by Cloud services. These could include solutions such as Natural Language Processing for Medical Records, and RPA with Insurance Claims.

Some vendors now have packaged solutions for the several of these horizontal and vertical market uses.

Financial Processes:

Many documents directly affect financial processes. Invoice Processing and AP and Receivables/Remittance Processing, for example, are often a focus as they are typically driven by the need to process incoming paper more effectively. An average invoice form contains around 120 characters of information that must be captured into an ERP system, typically SAP or Oracle Financials. Despite technology, much of this information is still keyed

manually. Using advanced capture capabilities to extract and validate data, then match it to the data in the ERP system eliminates the need for manual data entry, thereby reducing labor costs, and speeds invoice processing.

Lending/Mortgage Processing:

To date, most Capture Software solutions within Banking and Financial Services have been implemented in Retail and Branch banking, though Brokerages and Mutual Fund firms have also leveraged the technology. In this market, capture has traditionally consisted of scanning new wet signature paper-based applications for loans and mortgages, annuities, wealth management agreements, some signature cards and checks. Typically, it has been characterized by high-volume, batch scanning practices. Increasingly though, Banking is moving toward executing business processes closer to the point of transaction. Customers are challenging banks and other financial institutions to provide a faster response and a better overall experience. This is especially true in lending, where application forms and supporting documentation can be submitted online or via a mobile device and use digital signatures. By implementing recognition and classification technologies, perhaps even augmented with semantic understanding, organizations may reduce the cost of manually classifying mortgage applications and reduce the number of human touch points in the lending process.

Higher Education:

Though a more niche market, Higher Education has many document-centric processes that benefit from capture software and some of the advanced recognition capabilities it can provide. Transcript Processing is one such use case. Many universities and colleges are dealing with a growing population of students with existing course credits, who are transferring from other educational facilities or returning to school after being in another career or the military. Evaluating those credits in order to make admission decisions or accelerate course registration is often a time-consuming and largely manual process. For example, a university in the Southeastern US faced a backlog of transfer credit evaluations. Over the course of a decade, the university had experienced an enrollment increase of 10,000 students and had seen transfers increase by between 5 and 10% each year. The existing staff had not been expanded and the Registrar and Admissions offices were

reviewing more documents and transcripts using manual processes and legacy technology. Multiple people had to touch each document and manually key data into the university's PeopleSoft application. The university was unable to respond quickly to potential students and deliver admission decisions in a timely manner. Ultimately, this led to difficulties attracting transfer students. To address these challenges, the university implemented an advanced capture solution that did not rely on templates to do data extraction and that provided seamless integration with PeopleSoft. This solution automatically classifies incoming transcripts, extracts the necessary data and validates it with the business application. It enabled the university to significantly reduce the turnaround time on transcripts and deliver admissions decisions faster.

Conclusion

As the market evolves, Capture Software and Services continue to provide value to businesses of all sizes and government agencies. They enable any inbound data, regardless of its format, to be digitized, have metadata applied, and routed to appropriate personnel or directly into a process or application. Capture becomes the means of transforming multichannel data into information.

Consider how your organization currently makes uses of unstructured or semi-structured data. Look at the ways in which applying current and emerging technologies may improve your bottom line, enable faster customer or constituent service, or deliver competitive advantage. Opportunities remain for reducing operational costs and optimizing processes by taking paper out of the equation. New uses are emerging where capture and related capabilities can be leveraged even further. With capabilities such as auto-classification and auto-extraction, for example, capture becomes an automated process at the point of input. Machine learning and other AI techniques allow more data elements to be captured and interpreted quickly and accurately.

The next wave of capture technology will enable organizations to classify, interpret and understand incoming data from multiple channels and transform that data into information. That information will then feed business processes and support decision making, often in real time.

About the author:

Written by Karen Shegda for HSA: Ms. Shegda is a seasoned industry analyst and content creator with more than 35 years experience in consulting and writing about technologies, trends, vendors and products in the IT industry.   Formerly a Research Vice President at Gartner, Inc., Ms. Shegda has focused on content related technologies such as Enterprise Content Management (ECM), Imaging, Content Services, and Customer Communications Management (CCM) software. Ms. Shegda served on the Executive Board of the Association for Information and Image Management (AIIM), William Penn Chapter and holds a Bachelor of Arts degree in English.

About HSA, Inc. (Harvey Spencer Associates):
Since 1989, HSA, Inc. has specialized in electronic information capture technologies used to create, understand and extract meaningful information from semi structured and unstructured data to improve business process efficiencies. HSA assists organizations and vendors that support them with market support, insight and future direction.

About Hyland

Hyland is a leader in providing software solutions for managing content, processes and cases for organizations across the globe. For over 25 years, Hyland has enabled more than 19,000 organizations to digitalize their workplaces and fundamentally transform their operations. Named one of Fortune's Best Companies to Work For® since 2014, Hyland is widely known as both a great company to work for and a great company to do business with. For more information, please visit Hyland.com.