



BUSINESS EMPOWERMENT THROUGH ENTERPRISE CAPTURE

Capture as a standardized, shared enterprise service offers compelling business benefits



Table of Contents

Executive Summary.....	1
What is Enterprise Capture?.....	1
The Value of Enterprise Capture.....	2
The Evolution of Enterprise Capture.....	2
Key Elements of Enterprise Capture.....	4
Enterprise Capture Today.....	6
IBM Enters the Market.....	6
A Platform for Enterprise Capture.....	7
Conclusion.....	8

Executive Summary

In today's economic climate, every enterprise is challenged to increase the velocity of its business processes while operating in the most efficient manner possible. For the IT department, this means helping the business to grow revenue, increase cash flow, lower costs and improve customer service. IT is also called upon to streamline its operations through initiatives such as consolidating infrastructure and standardizing services.

Enterprise capture, which makes unstructured information understandable for input to business processes, is an emerging solution that addresses both challenges. It is estimated that 15 petabytes of new information are generated every day, a significant portion of which originates on paper, and some 80 percent of this information growth is unstructured data. Enterprise capture helps organizations leverage this information for better business results and more efficient operations.

Enterprise capture speeds the processing of unstructured information that arrives in media such as paper, email, fax, PDF and XML. By automatically grouping, understanding and extracting business-critical information from insurance claims, invoices, loan applications, tax returns, shipping manifests, sales orders and other transactional documents, enterprise capture can feed data to back-end systems at electronic speeds. This is far superior to labor intensive, costly and error prone processes for managing paper documents. And by consolidating multiple capture solutions into a single enterprise standard, enterprise capture can reduce the costs of managing infrastructure and vendors. In addition, capture can be integral to strategies for enterprise content management (ECM) and business process management (BPM).

The landscape of enterprise capture vendors has been changed in recent months by IBM's acquisition of Datacap, an ISV whose Taskmaster Capture solution is targeting a leadership position in enterprise capture. With the addition of this product to its line of content management solutions, IBM positions itself as a single-vendor provider of enterprise capture and ECM.

This white paper explores the potential of enterprise capture and explains the elements of a typical system. It also presents the views of the executives behind IBM Datacap Taskmaster Capture, giving their vision of the architecture, functionality and integration capabilities needed for a successful enterprise capture solution.

What is Enterprise Capture?

Enterprise capture uses common software to capture and extract information from documents enterprise-wide, either for a single business function in multiple locations, or for multiple business functions across the enterprise. Examples of this capability include a single capture solution for accounts payable/invoice processing implemented in multiple branches/countries, or a common solution to capture and process paper, email attachments and faxes serving claims processing, HR, accounts payable and orders/accounts receivable.

Enterprise capture requires an intelligent process that captures documents as they enter the enterprise. Document classification, data recognition and extraction, and routing are automatic. Typically, enterprise capture feeds extracted information to business processes via integration with case management, ERP, CRM, BI or other applications, while also indexing and preparing document images for ECM storage.

The Value of Enterprise Capture

Enterprise capture envisions making capture technology available throughout the enterprise as a standardized, shared service, replacing the multiple silos of capture and document management that typically dot the enterprise. This can offer both strategic and tactical business benefits, including the following:

- Increased business velocity from faster automated routing and processing of captured information. Benefits can accrue in any business process that relies on data from unstructured documents.
- Streamlined operations from “truncating” processes for managing paper by capturing documents as they enter the corporation
- Better compliance from immediate classification, indexing, addition of metadata and tracking of documents.
- Reduced security risk from eliminating the transport of paper documents to centralized centers or service bureaus for capture
- Cost savings from the reduced need for data entry, document transportation and document retention
- More efficient IT management and reduced costs from consolidated infrastructure and vendor management.

By enabling paper and other unstructured documents to be processed at digital speed, enterprise capture can fulfill the oft-stated promise of the paperless office.

The Evolution of Enterprise Capture

Enterprise capture has evolved from varied methods to capture information from business documents. From keying data from paper, to ad-hoc on-demand scanning in departments, to web-based scanning in the field, to high-speed batch scanning and forms processing, enterprise capture leverages advances in scanners, in recognition software for image classification and data extraction, and in the architecture and processing power of enterprise-class solutions.

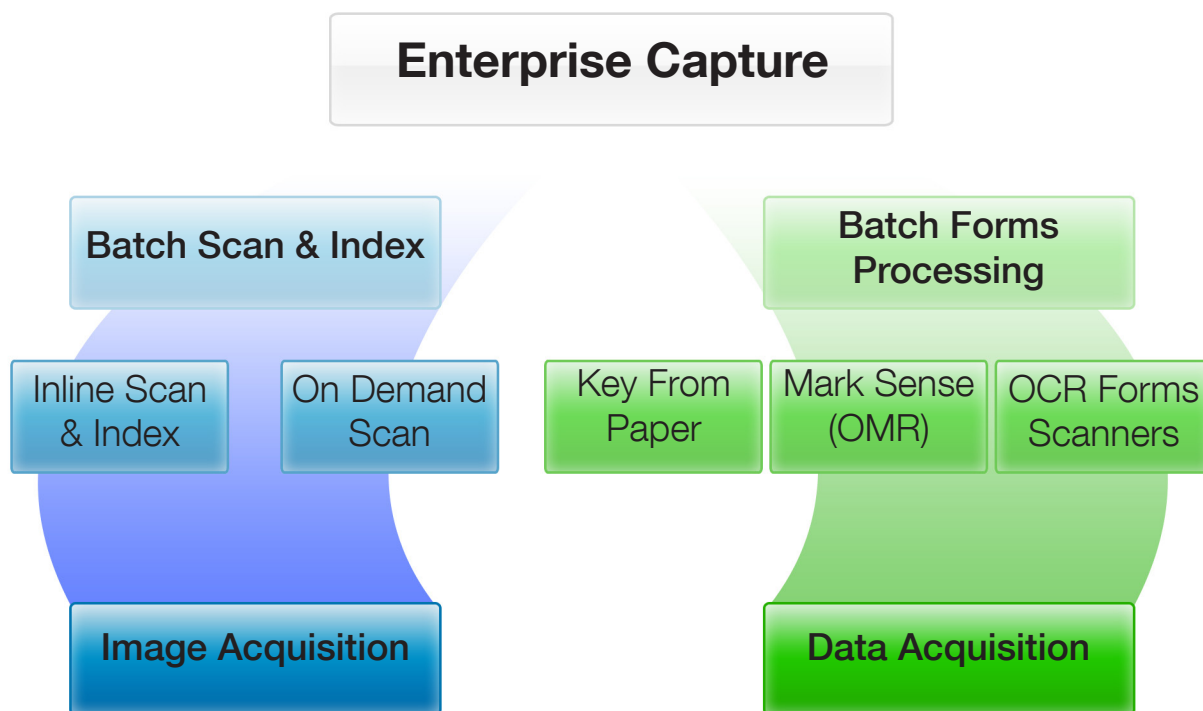


Figure 1 -Enterprise Capture brings together the two longstanding processes that companies still use in their document strategies. Image Acquisition comes from the microfilm industry, where documents are imaged and indexed for retrieval, while Data Acquisition is an outgrowth of the forms processing industry where data is extracted from documents for transaction purposes.

Historically, enterprise capture began in the mailrooms of Europe. Their mailroom software was developed to prepare and scan incoming mail and then electronically distribute it. This set the stage for enterprise capture, as it entails scanning and distributing multiple types of documents as they enter the enterprise. In the U.S., the U.S. Postal service has been precluded by law from offering an electronic delivery service, but key transactional-type mail, such as tax returns, forms and retail payments, has been pre-sorted through the use of post office boxes. It is then scanned and processed.

Enterprises may also use mailroom software to scan mail as it enters the organization. The capture software uses automated classification to decide on the routing. In Europe and elsewhere, this method has been expanded to help perform customer service and other business processes, a key requirement for enterprise capture.

Technological advances in capture dovetail nicely with market trends in enterprise software. Today's corporations realize business value from standardizing services across the enterprise. By consolidating IT platforms, vendors and processes, they reduce costs and simplify management. By making capture services available enterprise-wide, they can foster innovative solutions. As one or more departments achieve business value from the paperless office, others may follow. And a paperless solution for managing unstructured content can assist in compliance efforts.

Key Elements of Enterprise Capture

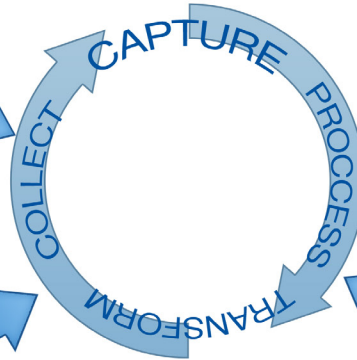
Capturing Documents at the Point of Entry

Increasingly, enterprises recognize the value of converting documents to images wherever and however they enter the corporation. Thus enterprise capture must support multiple ways of capturing documents at their point of entry.

Mailroom



Department



Field

Branch

Figure 2 -Enterprise Capture software must enable Image and Data Acquisition from every location where documents arrive in the organization, acting as a capture portal to funnel indexed images and accurate data from documents to all the target systems used by an organization, from ECM and ERP to BI and Analytics.

Central capture entails delivering documents to a mail branch, headquarters or post office box, where they are opened and scanned using a high-speed centralized scanner or small mailroom scanner(s). *Branch capture* enables received documents to be scanned as batches using a traditional scanner, or scanned on demand using a network scanner or MFP. Where agents or affiliates are involved, scanning occurs at their locations.

Departmental capture processes documents received via internal mail, or as a result of printing, by on-demand scanning in the department, usually with a shared network scanner or MFP. Departmental capture includes fax servers, relying on image-based processing to understand and extract data in the same way that scanned paper documents are processed. In addition, attachment files such as Word documents, PDFs and XML files, as well as the emails themselves, must be integrated into enterprise capture.

Field capture entails agents scanning documents immediately using smart mobile phones, tablets, mobile scanners or digital cameras, perhaps via a web interface. Examples included signed orders, checks, and receipts.

Intelligent Classification, Interrogation and Extraction

As documents are scanned, some will stay as indexed images (e.g., reports), while on others the data on the document is more important than the image (e.g., invoices). Enterprise capture starts to add value early through integration with business processes at time of capture.

The volume and variety of document types make it essential to minimize costly document prep, pre-sorting and data entry. Image and textual based classification software reduces this by identifying and grouping the images. This is followed by “interrogation” using a variety of pattern-based recognition technologies including OCR to understand the underlying data, and then validation and “extraction” to remove accurate data needed to feed business processes. Finally the images and/or data are selectively routed to the appropriate business areas, and to ECM systems for storage and management.

In order to accurately classify images, correctly understand them and ensure that valid and accurate data is extracted, the capture solution must integrate tightly with the appropriate business processes and enterprise systems. At all stages, metadata can be added to the images for better record control and management.

As an example of the fast time to value from enterprise capture, consider a bank customer who has applied for a mortgage. To gain approval, he needs to sit down with a loan officer and provide collateral materials, such as proof of income and financial statements. An enterprise capture solution enables the loan officer to scan the collateral while working with the customer, and the capture software automatically classifies and checks the documents, including immediate income and asset verification. Further, capture initiates a process where a bank employee analyzes the applicant's investments and suggests additional services such as asset management. In this way, enterprise capture helps the bank improve customer service and increase service revenue.

A Flexible, Extensible and Scalable Solution

Enterprise capture software must be flexible enough for easy adoption by different departments and processes throughout the organization. A true enterprise solution makes it easy to program the various rules and routines for document classification, data location and extraction required by the capture process. Once programmed, such rules ideally would be extensible, callable as web services by any application. As an example, customer account or social security lookup processes could be used by multiple applications outside of the formal capture process.

The high volume of documents in enterprise capture, and its distributed nature, suggest the need for a scalable and distributed architecture. Solutions should be able to leverage the power of today's multi-processor systems, and capture should be extensible to thin and thick clients via the web. Cloud computing, too, is important to enterprise capture. Cloud services enable enterprise capture to access remote features, functions and databases to better understand incoming documents, and to validate and verify extracted data. In addition, the cloud can provide the scalability, compute cycles and storage-on-demand required for enterprise capture deployments.

Enterprise Capture Today

Enterprise capture is in the early stages of adoption, with corporations installing enterprise capture in one of two ways. Some are taking an existing departmental system and expanding it to multiple departments. This does not necessarily change the method of operation into a true enterprise capture solution, but we believe it will evolve into this. Other users are purchasing a business-process-oriented solution, such as invoice processing integrated with procure-to-pay, and installing it in multiple sites, sometimes with remote scanning. This is more strategic, as it immediately delivers improved business value. In the future, we believe these two environments will combine to drive integrated capture solutions.

Note that enterprise capture solutions are likely to require business process reengineering to integrate capture with back-office business functions. While this raises the stakes for those considering enterprise capture, it also suggests the need for capture software that is flexible and easily customized to a range of business processes and applications.

IBM Enters the Market

Into this landscape comes IBM. On August 10, 2010, IBM announced the acquisition of Datacap Inc., a privately held provider of document capture software to organizations worldwide.

Datacap, founded in 1988, provides solutions that automate the capture, data extraction and classification of content for both structured and unstructured documents prior to storage in an ECM system. With an installed base of over 200 enterprises, Datacap offers industry specific applications for capture automation such as invoice processing, healthcare and government. Other capabilities include departmental or mid-market implementations for collaboration and basic content services; industry fast-start packages for customers and partners to build vertical solutions; and batch capture for many ECM repositories.

As part of IBM's ECM portfolio within IBM Software Solutions Group, Datacap's products enhance IBM's FileNet ECM offering, while adding support for non-IBM repositories and business processes. Taskmaster Capture, Datacap's flagship product, was designed for enterprise-wide capture deployments and positions IBM as a one-stop provider of comprehensive capture and ECM solutions.

The latest version of Datacap Taskmaster Capture, v. 8.01, has been integrated with all three of IBM's document image and management repositories: IBM FileNet Content Manager, IBM Content Manager and IBM FileNet Image Services. It also will provide intelligent document recognition for IBM FileNet Capture Professional. At the time of the acquisition, Ron Ercanbrack, then VP of Enterprise Content Management for IBM, said, "We've chosen to make Datacap's approach the foundation of IBM's document capture strategy. Datacap's approach to image capture, using sophisticated business rules management, sets it above the rest in the industry and provides the most complementary capabilities for IBM."

A Platform for Enterprise Capture

A key rationale for the IBM acquisition is the potential of Datacap Taskmaster Capture to drive enterprise capture deployments. Executives involved in developing the product and integrating it into IBM's ECM portfolio believe it meets key requirements for enterprise capture, which include the following:

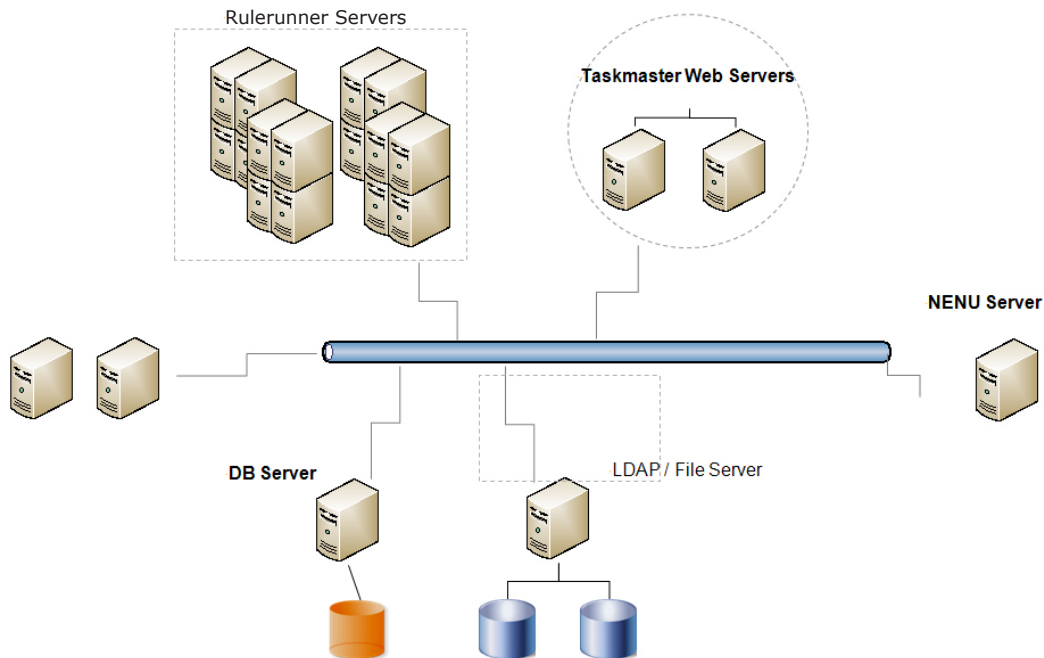


Figure 3 -This diagram shows Datacap Taskmaster Capture's scalability. Because the Rulerunner servers handle all the capture processing, such as image enhancement and recognition, as volumes rise, organizations simply add more Rulerunner servers to handle the additional work. As a result, Taskmaster can scale up to capture millions of documents per day.

Flexible Options for Capture and Processing

Datacap Taskmaster Capture features a three-tier, client-server architecture. Capture can be distributed via a thin-client web interface to the field, the branch and other points of document entry, while processing is delegated to thick clients and the server tier. As an organization grows or as more documents and departments are added, processing power can be added to maintain performance levels.

Adaptable, Extensible Process Management

For capture to become an enterprise standard, automated capture processes must be easy to program and extend across the organization. The engine for this in Datacap Taskmaster is Rulerunner Service, a configurable procedural rules engine that drives virtually all document capture processing functions, including scanning, document ID, image enhancement, text recognition and formatting data for export. Hundreds of pre-built rules and actions are included as discrete executable functions stored in a library. Using Datacap Studio, an application building tool, no programming experience is required to create or modify rules. Rules developed for one application can easily be applied to other applications and business processes.

Scalable Capture Callable Across the Enterprise

Enterprise-class capture solutions require scalable processing power, as well as capture processes and routines to be callable as a service. Datacap Rulerunner Enterprise provides scalability by enabling virtualized instances of Rulerunner Service to run on multi-core processors. In addition, Datacap Taskmaster Capture supports Services Oriented Architecture (SOA) to enable capture processes such as document identification and data location features to be callable outside of formal capture solutions.

Conclusion

Enterprise capture can improve an organization's business results by streamlining processes that involve images, paper and other unstructured documents. Keys to achieving such benefits include choosing capture software that is flexible, adaptable and scalable across the enterprise, and then replacing multiple silos of capture with an enterprise standard. This will enable multiple departments to integrate capture into a variety of business functions. Although enterprise capture is still early in its maturity cycle, the potential makes it worthwhile for every enterprise to explore.

Empowering the Organization with Enterprise Capture

An interview with Scott Blau, Datacap's co-founder and director of enterprise content management capture solutions, IBM; and Reggie Twigg, global product manager, ECM content and capture integration, IBM.

Q. What is the relationship between capture and ECM?

Reggie Twigg: IBM has been selling ECM systems to enterprises for some time. An ECM system can be considered the system of record for unstructured content, much like an ERP system is the system of record for financial and supply chain data. I think we're going to start seeing organizations standardizing on capture for systems of record.

Q. What capabilities do customers need to make enterprise capture successful?

Scott Blau: At the enterprise level, customers need to be empowered by this technology; to be able take control if it, use it wherever they need it and tailor it to their requirements. It's far from a one-size-fits-all solution, because every deployment has unique nuances. The theme we've worked on for the past five years is empowerment by reducing the amount of setup, and exposing a level of control so that business users can easily orchestrate individual capture steps to their exact needs.

Q. What's makes enterprise capture different from traditional capture solutions?

Reggie Twigg: It's not enterprise capture unless there is an enterprise standard -- all departments should do capture on one technology and follow certain best practices. We typically see islands of capture throughout the enterprise, different technologies bought by different groups that aren't integrated. There's no consistency and they're dealing with 4-5-6 or more vendors. This should cause the CIO to ask, why do we need all these vendors that are providing the same technology?

Q. How can enterprise capture empower an organization?

Scott Blau: Traditionally, capture has been viewed as a silo, where you plug this capability into the front end of document management and that's the only place you use it. You're doing OCR, extracting information, integrating with the email and fax system, all on a somewhat rigid front end. But say follow-on documents come in via email to a group that is managing a fraud case. You could print them out and send them over to the department doing formal capture, but that's clunky. Enterprise capture is calling the same capability you have in production capture into the case management system, and getting the same result. Being able to leverage existing technology greatly speeds time to market and lowers development and ownership costs.

About Harvey Spencer & Associates

Since 1989, Harvey Spencer Associates based in New York (US tel: 1-631-368-8393) has been specializing in electronic information (image based and electronic transaction) capture technologies. Our services include Market Analysis, Technology Planning Assistance, Product Positioning, Product Management, Client Sponsored Research and Strategic Planning Services. these products include high speed document scanning hardware, image acquisition software, character recognition software (OCR, ICR), optical mark recognition, barcode recognition and other pattern recognition and classification tools.

As we move more to an ebusiness environment, it is clear that paper based information must be managed more effectively. The best way to do this is to convert the paper to usable electronic representations as fast as possible. We are specialists in the hardware and software used for capturing images and data from paper or microfilm and in electronic document imaging and forms processing software with OCR and other associated recognition technologies. We believe that advanced pattern recognition is critical to creating understanding of these documents and that XML and related technologies are fundamental to the interchange of these documents.

We are the only specialist company following this market and to assist in adding value in these areas. We advise and take a leadership role in developing standards and other technologies. As an analyst company we have measured and analyzed the market for document capture software.

For more information, please contact: jim.everett@hsassocs.com

About, Co-Sponsor, IBM Enterprise Content Management

IBM Enterprise Content Management software enables the world's top companies to make better decisions, faster. By gaining control of unstructured information, companies can access, collaborate and influence business decisions in new ways, making content a first-class source of insight. With industry-specific IBM ECM solutions, companies can capture, manage and share content throughout its lifecycle to help ensure compliance, reduce costs and maximize productivity. The IBM ECM portfolio includes a wide array of capabilities that integrate with existing systems to help organizations maximize the value of information including: document capture and imaging; social content management; advanced case management; information lifecycle governance and content analytics. More than 13,000 global companies, organizations and government organizations rely on IBM ECM to improve performance and remain competitive through innovation.

www.ibm.com/software/data/content-management/