





Foreword

For you as a business professional, time is a precious commodity. You constantly need to distill concepts, evaluate options, and manage complex transactions. When you need information, you need it in a form that can be assimilated quickly—forget the mind-numbing detail and get to the point.

With that in mind, we've developed our series of 15-minute guides to essential topics in the area of information management. This particular guide focuses on the importance of recovery management and how it can help make your business more resilient and successful.

In about 15 minutes, we'll define document and image processing, explain why it's important, outline the benefits, suggest a deployment framework, and provide resources for you to learn more. We think you'll agree that it will be 15 minutes well spent.

Table of Contents

<i>Introduction: The problem with paper</i>	4
<i>How to win the paper chase: Automating the capture of incoming documents</i>	5
<i>Integrating transactional content with enterprise workflow</i>	5
<i>The case for an end-to-end document and image processing solution</i>	6
<i>The steps of managing transactional content</i>	7
Step 1: Capture	7
Step 2: Integration with workflow	8
Step 3: Access with appropriate controls	9
Step 4: Archiving and storing	9
<i>Managing transactional content for compliance</i>	10
<i>Ready to get started?</i>	11





Introduction: The problem with paper

For centuries, business has run on paper. In paper-intensive industries, that's still the case. Consider mortgage brokers, for instance. A typical application file for a customer loan might include hundreds of pages of electronic forms, copies of written documents and original certificates of title, digital pictures, faxes, and even handwritten notes.

But as the pace of business has accelerated, paper has become a burden and a bottleneck. It slows productivity because only one person can handle a paper document at a time (unless it is copied, which takes more time and further erodes efficiency). In an increasingly stringent regulatory environment, paper is also a legal risk. Having to search for content in response to a legal demand could lead to costly penalties and lost productivity if a document is lost or missing.

To avoid these and other problems, organizations are streamlining manual, time-consuming processes. The newest trend is the use of digitization and other capture technologies (such as document classification engines and optical character recognition) to integrate information into workflow at the very beginning of the process, as soon as it enters the system. Designed for paper-intensive applications such as financial services, insurance, healthcare, and accounts payable, these technologies are engineered to automatically convert large volumes of paper documents and other original content to digital form and make them available for business processes.

Transactional content: Term coined by Forrester Research to describe “content [which typically] originates outside an organization from external parties—customers or partners—and relies on workflow or business process management (BPM) to drive transactional, back-office business processes. In some cases, the content not only triggers internal processes, but is the basis for the transaction itself.”

Document and image processing: A software solution that enables organizations to capture, process, and access electronic images of documents. Content may include paper and electronic documents, photos, reports, computer-generated reports, XML data, and electronic forms.

According to Boston's Delphi Group:

- 15 percent of all paper handled in a typical business gets lost
- 30 percent of all employees' time is spent looking for documents
- Each lost piece of paper costs a business \$120

The benefits of getting rid of paper can be immediate and dramatic: reduced cycle time, improved efficiency, lower costs, and better compliance simply by automating manual processes. However, an organization that wants to get maximum leverage from its document and image processing strategy will consider a holistic, end-to-end solution, one in which transactional content is managed in a central repository along with all the other content necessary to conduct business.

Organizations that excel in management of content at every point will get their work done faster—and often win business away from less agile competitors. They'll make better decisions with an integrated view of all relevant information. They'll improve customer service with better, faster, more comprehensive response to inquiries. They'll avoid the regulatory cost and embarrassment that happens when a document can't be located in response to a litigation request. And—based on the typical experience of EMC customers—they'll enjoy a 30 percent to 50 percent gain in overall cost and process efficiencies.

In this 15 Minute Guide, we'll discuss major strategies for leveraging transactional content through document and image processing solutions—including the benefits of automating information input, the value of integrating content with business workflow, and the case for an end-to-end solution for making all content available throughout your enterprise on a unified platform.

We'll examine the steps of managing transactional content: capturing, processing, accessing, and archiving. We'll look at its use for regulatory compliance. We'll share real-world stories of companies that are successfully managing transactional content today. And we'll talk about the options to get started in your own organization when you're ready.

How to win the paper chase: Automating the capture of incoming documents

Simply by implementing an input management solution, such as EMC Captiva® InputAccel™, an organization can achieve significant payback right away. The organization will save time because data enters the workflow far faster than with manual processing. It will reduce costs because less human interaction is needed—a huge consideration in any organization where the volume of information and documents to be processed continues to swell, but headcount is fixed for budget reasons. It will improve the quality of information by avoiding the mistakes that come with manual data entry.

In addition, by using intelligent document recognition software, such as EMC Captiva Dispatcher, organizations can automatically identify documents as they are captured and route them to the appropriate organization, individual, or application. A high-speed solution for automatic classification, indexing, extraction, and routing helps businesses handle the vast amount of incoming information they receive on a daily basis in a fast and efficient manner.

In some instances, this efficiency of processing might translate directly into a competitive advantage. In the loan origination business, for example, loan applications can include hundreds of pages. By digitally capturing loan documents, extracting key data from those documents, and making that data simultaneously accessible to all individuals in loan processing, a lender can process more loans more accurately and quickly.

Integrating transactional content with enterprise workflow

A century ago, Henry Ford proved it was much faster and cheaper to manufacture cars on an assembly line than to build them individually from beginning to end. Instead of completing one product, then starting another, an automated system provides efficiency gains by fine-tuning the repetitive processes. It makes some steps more efficient while eliminating other, manual steps.

Business process management (BPM) brings this manufacturing concept to software applications such as claims processing and new account set-ups where customers complete an online form for streamlined processing. Companies that are highly successful today generally employ BPM in many of their operations. Companies that don't employ these processes risk mounting costs and non-compliance with legal requirements.

How does transactional content fit into BPM? Many of the businesses where transactional content can be found have highly repetitive processes that are good candidates for employing BPM. Think of loan origination, invoice processing, claims processing, case management, and new account setup. The key is providing a way for the business applications to “know” what is in the documents used in these businesses so the information they contain can be leveraged with other automated processes. This is done through classifying content and applying metadata—data that describes the data—at the time the document enters the workflow.

Once transactional content is made available to workflow, an organization can look forward to:

- Efficient reuse of standard content, such as a loan application which becomes a template for generating new application documents
- Better decisions, because more relevant information is available for loan scoring or other business processes
- Reduced risk through standardized templates which ensure that processes are in compliance with legal and business rules
- Improved agility through increased responsiveness to both predictable and unexpected business or competitive factors

These benefits are inherent with highly automated processes. But document and imaging processing solutions extend the gains in efficiency and control to far more types of content, providing a much richer information resource to drive the business.

Unified platform saves money, improves competitive position at Cincinnati Insurance

The Cincinnati Insurance Company markets property casualty insurance in 32 states, with a business model focused on keeping costs and premiums low. In order to reduce its claim file administration time, Cincinnati Insurance replaced its paper-based claims processing with an automated system, including a central document repository.

The new system has reduced the claims entry time from 18-30 days to one day, while providing users with sub-second access to claims files in order to assure better customer service. The new system also reduces legal risk by storing all data in an unalterable, secure state. In addition, Cincinnati Insurance was able to lower costs by storing more data on EMC Centera™ content addressed storage devices.

Says Gary Givler, AVP of the Headquarters Claims department, “an integrated, virtual claim file is critical to our success in today’s increasingly competitive industry. With our new claims management system, we document more claims with fewer people. We also make better and more informed decisions and can respond quickly because our information is current and centrally available. In turn, our costs have decreased and customer satisfaction has improved.”

The case for an end-to-end document and image processing solution

An auto accident insurance claim is an example of a situation where many different people, inside and outside the insurance company, need access to the same content: police and accident investigators, body shops and estimators, rental car companies, litigators, internal review staff, and of course the policyholders and agents.

In a “paper world,” such a claim might be held up because the materials required were in a physical file jacket that was moving from desk to desk. One department couldn’t work with it until another was finished. Yet even companies that have succeeded in digitizing transactional content may not get all the benefits of content because they work within multiple and incompatible applications or because information is housed in separate “silos” of content.

Many organizations have applied first-generation “point solutions” to handle specific functions—one system for accounting, another for materials handling, and so on. But problems arise when systems hand off data from one to the next: if the materials-handling application couldn’t talk to accounting, then information would need to be re-entered or translated. Similarly, a user who wanted to cite an invoice or other document in a correspondence might have to attach a hard copy because he or she could not access the document in the application where it was created.

All of these problems have become magnified in recent years through a wave of consolidation and mergers and acquisitions. A newly merged company might have two different—and incompatible—applications doing the same thing.

This is a key reason companies are turning to end-to-end solutions from companies such as EMC, which leverages a single repository to handle all content and make it available to any application that needs it. The content is also managed so the organization knows who has it, where it came from, and how it has been revised—a perspective that facilitates business clarity and is essential for regulatory compliance.

An end-to-end solution also enables the management of content throughout its lifecycle—recognizing that not all information is used in the same way, or has the same value, at all times. High value information, such as active customer account files, will be kept on local, high performance storage devices. Inactive records can be moved to less-accessible devices or deleted.

End-to-end solution increases efficiency at Commerzbank

In the competitive retail and commercial banking marketplace, improving operational efficiency has become essential for survival. That's why Commerzbank, one of Europe's leading private sector banks with consolidated assets of nearly 400 billion euros, chose to streamline its corporate credit lending process using EMC Documentum® content management software.

Business rules are now modeled within a new automated system and all documentation (including the loan application form, risk analysis, return on investment spreadsheets, presentations, correspondence, annual reports, meeting agendas, and minutes) is stored electronically in a central repository. Client-generated correspondence and other documents originating in paper form are acquired by scanning. Loan applications, meeting minutes, and all supporting credit documentation are entered using XML-based online forms.

When the system was first implemented, Commerzbank had some 100,000 pages of documentation related to the previous year's business, growing at a rate of approximately 150 pages per day (or one file cabinet per month). Today, the need to retain paper has been virtually eliminated. With a central repository, groups can work in tandem on a credit application—resulting in greater efficiency and faster time to final approval. Another benefit: the bank's internal audit rating has dramatically improved due to better business policies for capturing an audit trail and for avoiding the potential for lost or misplaced documents.

The steps of managing transactional content

Now let's take a closer look at the process: what actually happens as paper and other transactional content is transformed into digital documents and data and delivered into enterprise workflow, where it can be utilized, and how it's ultimately archived or deleted.

Step 1: Capture

The more customer interactions and supporting documentation can be captured and stored electronically in a single repository, the more valuable the content becomes. The original records may be:

- Captured scans from high-volume imaging solutions, received as paper or fax and translated via optical character readers (OCR) or saved with appropriate information
- Electronic forms such as new account setup applications submitted online
- System-generated content captured through computer output to laser disk applications (COLD) or enterprise report managers (ERM)
- Documents created in authoring applications such as Microsoft Office

In general, incoming content will go through some or all of the following procedures:

- **Capture:** Scanned images may be refined through de-speckling, page orientation, border removal, de-skewing, and other image pre-processing so that “noise” and irrelevant material can be removed or ignored and only transactional content retained.
- **Classification:** A solution such as EMC Captiva utilizes text, images, and knowledge-based analysis techniques to automatically identify documents and prepare them for processing—without the need for manual sorting or separator sheets. Such analysis identifies various document types, such as an appraisal form or a W-2, for example, and enables these documents to be automatically routed in accordance with business rules. Documents are then indexed or “tagged” so they can be easily retrieved from an enterprise content management repository or other system.
- **Extraction:** In many cases, data from transactional documents needs to be extracted and made available to users in a variety of other applications. For example, a handwritten new account application might be optically recognized and translated to machine readable text.
- **Validation:** Once it has been extracted, data can be validated against line-of-business applications to ensure its accuracy and integrity. As an example, a new account application could be compared to business rules to ensure all fields have been filled out.
- **Export:** During the export process, the data is made available in the appropriate format—such as PDF, XML, or file system—and converted directly for use in databases such as Microsoft, Oracle, Informix, and Sybase. For companies that use an EMC Documentum content management system, all data and documents can also be exported into the Documentum repository.

Step 2: Integration with workflow

Once documents have been captured in a repository, they can be put to work. We’ve already discussed the benefits of automated workflow to dramatically speed throughput, improve customer service, and avoid errors. The workflow system allows automation of human interaction. Steps may include:

- Applying business rules via BPM services to automate repetitive tasks that have traditionally been manual processes
- Integration with line of business (LOB) applications such as ERP, CRM, and Microsoft Office, so additional information is made available for correspondence, research, and collaboration
- Enforcement of records management and retention policies to determine what documents are kept, under what security measures, and for how long
- Merging acquired content with appropriate document output management templates to create highly personalized communications, such as customer service letters

In some cases acquired transactional content is merged into “virtual files,” which contain all relevant documents and content for a single case or customer. Virtual files can be efficiently routed and organized so nothing is lost or ignored. And queue management may be used to automatically route information to the right people so responses to inquiries can be prompt.

In addition, rules management may be used to automate as many processes as possible and to minimize the need for manual intervention. For example, a company may apply a rule to an invoice that is under \$5,000, enabling it to be automatically approved and paid without any human interaction.

Step 3. Access with appropriate controls

Managing business files in digital form gives employees the global, on-demand access to content that is necessary to ensure a 360-degree view of customer data and business transactions.

Once documents are digitized and stored in a single repository, critical data can be easily extracted to quickly access and intelligently process a customer's request. Employees can search, view, and annotate documents. Information retrieval occurs more quickly with powerful metadata search capabilities or by browsing through the folders and categories that organize the repository.

Access also facilitates collaboration to improve the speed and quality of decisions and service. Perhaps the claims management team will assemble for a few minutes each day to review and confirm pending activity—a process that would be far more cumbersome under a paper-based system. Collaboration can also be triggered automatically as part of business processes; for example, officers might be summoned by e-mail to a “virtual conference room” (such as EMC Documentum eRoom®) to review certain claims based on criteria defined in business rules.

In customer service, better access to a holistic view of information leads to better customer satisfaction and fewer lost customers due to service problems. It also saves time and money because more issues can be resolved on the initial contact.

Enabling greater access delivers benefits, but access must also be controlled to protect customer confidentiality and assure the integrity of records. Typical controls include:

- Security measures to determine who can access each document and how often
- Annotation capabilities that include information about who commented on documents and when

Step 4: Archiving and storing

Generally there is a point in content processing where a document is no longer active. At that point it can be “logically laminated” and placed under formal records management control. It can be checked in and out but, if it is altered in any way, a new record will be made including metadata that describes who made the change and when. This ensures the record is protected in compliance with regulatory and business rules. It also sets into play any retention policies as to how long it is stored and eventual disposition.

The choice of a storage strategy is important with transactional content because of the sheer volume of documents. An organization can realize significant savings by moving archived documents from “primary,” high-performance storage devices (such as local servers) to “secondary” devices (such as EMC Centera content addressed storage) for less frequent access. The key requirements in storing this content are:

- **Authenticity:** The validity of documents must be protected with appropriate security measures which control access and limit the ability to revise or delete. The system must also provide an audit trail to show everything that happens to documents in storage.
- **Longevity:** The system needs to have built-in protection against media failure, system shutdowns, or external threat. Typically data is automatically replicated at secondary locations, and self-healing features continually monitor data integrity and make repairs as needed.
- **Accessibility:** Regulatory compliance (“e-discovery”) might involve the rapid recovery of hundreds or thousands of archived records in response to a subpoena or other legal request. In such a scenario, the ability to respond quickly might have an impact on litigation costs and help companies avoid penalties for non-compliance.

Content-addressed storage (CAS) is a relatively new approach to meeting these storage requirements. Content addressing eliminates the need for applications to understand and manage the physical location of information on storage media. Instead, addresses are calculated based on the content itself and serve as a “virtual claim check” that applications use to find and retrieve stored objects. This not only simplifies the task of managing huge numbers of objects, but also provides a digital fingerprint for the content, ensuring its authenticity. CAS also provides self-management and self-configuration capabilities to allow organizations to manage more content with fewer staff at a dramatically lower cost than traditional archiving management methods.

Automating the deal lifecycle at Boston Capital

A pioneer in the field of affordable housing, Boston Capital underwrites and acquires 80 to 100 multifamily properties each year. The process of putting these deals together is highly transactional and generates large volumes of paper, with individual documents that include hundreds of pages.

According to Brian Madden, director of information technology, “Our direct costs for onsite and offsite document storage as well as the administrative costs for filing and photocopying all of that paper were growing rapidly. And our indirect costs were also increasing proportionally—from the time it took to search for a document to the time wasted working on incorrect versions.”

To boost productivity and profitability, Boston Capital implemented a central repository for managing its deal documents based on EMC Documentum content management integrated with EMC Captiva InputAccel for document acquisition. The new system:

- Provides a single access point for all content necessitated by a deal
- Streamlines search and retrieval
- Allows users to electronically route documents into the electronic workflow

Today, says Madden, “acquisition and asset management teams spend less time searching for documents and more time managing their deals. Likewise, our tax department has significantly reduced the amount of paperwork that accumulates each tax season; they now review tax returns via an electronic workflow in conjunction with our external CPA firms. There’s no question that this platform has become an integral part of our business, and is essential to our day-to-day operations.”

Managing transactional content for compliance

New regulations, ranging from Sarbanes-Oxley to HIPAA, place increasing pressure on organizations to store content. This ensures the content’s integrity and enables it to be readily located and produced in response to subpoenas and litigation requests. There are many scenarios, ranging from a policyholder’s lawsuit to a regulatory body’s investigation, in which an organization might need to produce transactional content on demand. The control and accessibility of an end-to-end document and image processing system can save significant research costs and potential penalties while providing peace of mind. Such a system can:

- Eliminate the inherent risk of paper documents which are easily lost, misplaced, altered, or exposed inappropriately (Under HIPAA, for example, simply leaving a patient record unsecured on a reception counter can be a violation subject to significant fines).
- Ensure the integrity of documents—so you can authenticate that the records have not been moved or altered.

- Provide a verifiable audit trail that shows when documents were created and revised and who accessed them.
- Set and enforce policies for safe disposition of documents when their retention period has ended. This is a valuable capability because any existing record can potentially be subpoenaed in a court case—even if it is obsolete and its content is unknown to you.
- Enable fast and cost-efficient access to documents in the event of litigation.

Ready to get started?

If you'd like to start enjoying the benefits of managing transactional content with document and imaging processing, here are several places where you might begin:

Like many companies, you might start with the simple objective of reducing the amount of paper you deal with. Many high-volume capture solutions (including Captiva InputAccel) are designed to do exactly that. By simply digitizing transactional content, you can improve efficiency while eliminating the security concerns that go with hardcopy documents. One benefit of InputAccel is that it allows organizations to capture content and archive it to the Documentum repository, which itself provides a number of benefits, including the ability to ensure regulatory compliance.

Or, you might approach a specific departmental or functional area. If most of your manual processing occurs in accounts payable, for example, an automated invoice processing system might yield faster benefits in less time and at a lower cost than if you were to automate the entire organization. Take care, however, to ensure that this point solution can be integrated with other applications later on and will not become permanently siloed.

Or, you might start from the outset with the objective of embracing an enterprise platform which allows you to buy the pieces you need and then integrate other content management solutions—such as web content management, digital asset management, records management, and others—if and when you need them.

That's the path taken by Cincinnati Insurance, mentioned earlier. This organization started with the Documentum enterprise document management solution for Department of Insurance (DOI) submissions, then gradually added document capture, workflow processing, SAP® integration, dynamic content publishing and enterprise report management—all Documentum and Captiva products—as well as EMC Centera storage for its benefits of ROI and compliance.

But whatever approach you take, your effort to transform transactional content should bring significant benefits including:

- Managing all types of business information, from high-volume scanned documents to digital pictures, application documents, e-mail, and more, in a common format with validation of content
- Eliminating the inefficiency, expense, and risk of a paper environment in which documents are handled sequentially and can be easily misplaced or lost
- Providing overall control for business processes as required by your operations, industry standards, or regulatory factors
- Improving customer service with better, faster, more complete response
- Enhancing your market position relative to competitors who still struggle with manual processes and paper documents



About EMC

EMC Corporation (NYSE: EMC) is the world leader in products, services, and solutions for information management and storage that help organizations extract the maximum value from their information, at the lowest total cost, across every point in the information lifecycle. Information about EMC's products and services can be found at www.EMC.com

To learn more about document and image processing from EMC, visit software.emc.com or call **800.607.9546** (outside the U.S.: +1.925.600.5802).



EMC Corporation
176 South Street
Hopkinton, MA 01748
1-508-435-1000
In North America 1-866-464-7381
www.EMC.com

EMC, EMC, and where information lives are registered trademarks of EMC Corporation. Captiva, Documentum, and eRoom are registered trademarks and Centera and InputAccel are trademarks of EMC Corporation. SAP is a registered trademark of SAP AG in Germany and in several other countries. All other trademarks used herein are the property of their respective owners.

© 2006 EMC Corporation. All rights reserved. Produced in the USA. 5/06

S10070506V1