



# EMC Captiva FormWare

Enterprise-level  
forms processing



## The Big Picture:

- EMC Captiva's award-winning FormWare automatically identifies, extracts and perfects data from virtually any type of form and exports this data into back-end databases or systems including document or content management systems. FormWare customers speed the form processing process, improve information accuracy and reduce operational costs.

## Automating Forms Processing

FormWare by EMC Captiva is an essential tool for businesses processing large volumes of structured forms, including insurance claims, credit applications, tax forms and any other type of standard business form. Regardless of the source - paper, fax or electronic data streams - FormWare can collect, perfect and export the data from forms to any database, workflow, document and content management system without the need for manual data entry.

## Out-of-the Box Functionality, Enterprise-wide Scalability

FormWare is a modular solution comprised of various building blocks. The system is 100% scalable and can be deployed at a departmental level and scaled easily to an enterprise-wide solution when processing demands reach hundreds of thousands of images per day.

A typical FormWare workflow could contain the following modules:

### Capture

FormWare acts as a virtual "front door" to enterprise applications. This module also allows users to import images and electronic data streams from other sources, including unattended scanners, fax servers and the Web. Single forms or batches can be scanned, flexible naming conventions can be applied, and screen displays may be customized to provide a simple interface for scanner operators. A rescan module also allows users to easily rescan single forms that have been rejected at a later stage in the process without disrupting the overall workflow.

### Clean-Up:

Once documents have been imported or scanned into the system, FormWare implements a rigorous, fully automated image cleanup process. Image cleanup may involve despeckling, "noise reduction," deskewing, page orientation and border removal, as well as fourteen other image pre-processing functions.

### Classification:

FormWare utilizes EMC Captiva's leading page and document classification technology to identify document types.

### Data Extraction:

FormWare uses multiple recognition technologies to automatically extract data from documents, beginning with the very first page received.

### Validate:

Extracted data is validated against existing data contained in ERP, ECM or other information repositories.

### Export:

The export module outputs data and images in many formats, such as XML, PDF, HTML, ASCII and Word. Interfaces to ODBC-compliant databases and enterprise content management systems and storage and retrieval applications are available, and custom export formats can be easily implemented.

### Image Quality Control:

FormWare provides both unattended image quality checks and operator-attended visual inspection capabilities to ensure that optimal images are always captured. Inferior images are detected and routed back to rescan.

## In a Class by Itself

One of FormWare's greatest strengths is its ability to identify heterogeneous document types using sophisticated classification technology. The technology works by automatically learning the attributes of existing documents and using them as a basis for how to classify new incoming documents that arrive throughout the various departments of an enterprise. For example, the classification engine can perform full-page classification that takes a macroscopic view of a document in order to make intelligent decisions about what type of document is being classified or it can "zoom-in" on the proximity of word anchors on the page.

## Easy Setup and Administration

With FormWare, system setup, administration and maintenance are straightforward and include many options. EMC Captiva provides tools that enable flexible setup to create a truly enterprise-specific solution without the need for highly skilled programming resources.

## Administration Console

FormWare's Administration Console provides system administrators with a detailed record of all unprocessed, in-process and completed work. The system tracks detailed performance data for scanners, recognition engines and data entry operators. System administrators can view a summary of all jobs in progress and use that information to prioritize work or change system configurations to meet critical deadlines.

## True Parallel Processing

FormWare's parallel processing capabilities allow multiple clients to work on the same batch at the same time to achieve maximum efficiency. Documents themselves can be virtually disassembled to allow processing tasks to be delegate simultaneously to specific workstations before being routed back to the system for re-assembly and completion.

## FormWare in Action

One of the most noteworthy applications of FormWare is its use by tax authorities. FormWare is used by the revenue departments in numerous states to capture and manage the information contained on tax forms in an accurate, timely and cost-effective manner. Once the tax forms have been scanned, FormWare captures the original document, "reads" the date on the form via optical character recognition (OCR) technology, extracts the important data and exports it to the system. The result is a decrease in the costs of processing, storing and retrieving information, and a reduction in cycle time for paper and electronic filing, as well as for remittance processing and funding.

