

XenDocs Content Capture

Background

Content often starts in the form of a paper document that must be scanned, but also includes native electronic files, such as documents received via email or from a fax server.

“Content capture” is the process of gathering, classifying and indexing content (paper and e-documents) from various client applications and external systems for use within XenDocs™.

Powering Your Paperless Solutions

With XenDocs, content can be captured using specialized third party content capture applications or with native content capture tools within XenDocs. In either case, the process can be highly automated using a range of tools, depending on the type of document.

While documents must be “classified” before they are saved to XenDocs, the indexing process – where searchable metadata is associated with the document – can occur either before or after content is saved to the XenDocs Content Server.

For paper documents that are scanned into the system, automated recognition tools can be used to “read” information from the documents. This data can be used as searchable metadata within XenDocs or exported to an external line of business application or database.

XenDocs Capture Toolset

The XenDocs Capture Toolset provides a wide range of tools for connecting external applications to your XenDocs Content Server.

XenDocs Scan Module

XenDocs Scan is easy to use and tightly integrated with the XenDocs Content Server. It is ideal for users with occasional or distributed scanning requirements.

Third Party Application Adapters

XenDocs includes certified adapters for the leading production-oriented document and data capture applications, including:

- Abbyy FormReader/FlexiCapture
- Cardiff TeleForm
- Cardiff LiquidOffice
- Kofax Ascent Capture

XenDocs Case Study Overview

County of San Diego

AP Processing with Oracle Financials

Problem: Invoice documents that are received at remote departmental locations for processing were routed via inter-county mail for approval by auditors. Manual invoice processing led to excessive manual labor costs, long processing times and lost documents.

Solution: Using Fujitsu scanners and Ascent Capture from Kofax Software, XenDocs was integrated with the County of San Diego’s Oracle Financials ERP system and deployed throughout all 58 departments. The system provides a workflow process for AP staff working within each department to scan and process invoices remotely with minimal data entry. XenDocs then enables auditors to centrally review and approve invoices for payment.

Results: Enormous labor savings due to elimination of paper handling to process AP vendor invoices, while also enabling the County to capture more early payment discounts from vendors.

“Each invoice is reviewed three to five times during the AP invoice lifecycle. Previously, our AP staff spent anywhere from several minutes to several days locating the necessary paperwork each time. Now they can locate and view each invoice within a few seconds.”

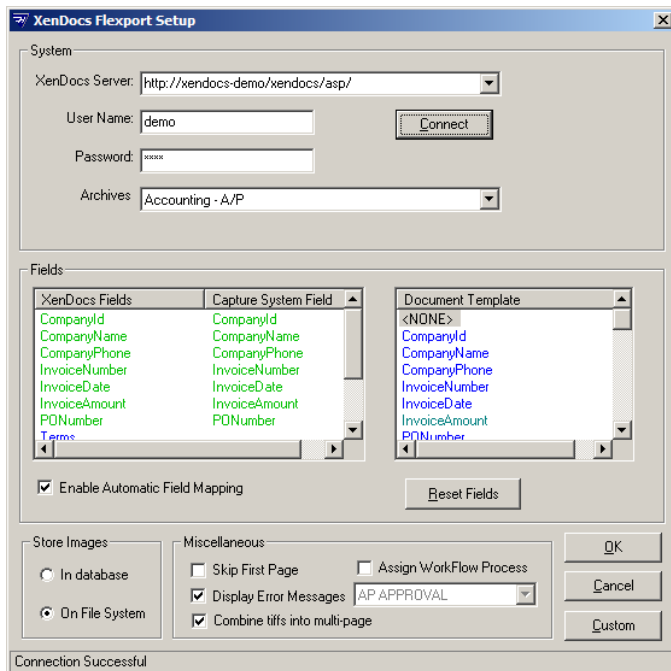
**- Carla Jane Mann
Procure-to-Pay Project Lead
County of San Diego**

XenDocs Import Service

Electronic files can be easily classified and imported using the XenDocs Import Service. This module allows you to define one or more “import folders” on your network, including a configurable profile for each folder. Based on the folder’s profile, the system imports the documents into the specified XenDocs collection. As the content is imported, each document is automatically indexed based on file attributes and/or data in a text or XML file.

FlexPort Setup Component

This graphical component provides a simple interface allowing users to easily map fields from various 3rd party applications to related document collections in their XenDocs Content Server.



Capture SDK / Toolkit

If your needs extend beyond the off-the-shelf capture toolset, then customized content capture solutions can be developed with the XenDocs Capture SDK. This developer toolkit is the foundation for all of the components in the XenDocs Capture Toolset.

The XenDocs Capture SDK is based on a leading-edge Service Oriented Architecture (SOA) and the Windows Communication Framework (WCF). This design allows any XenDocs-enabled client application to run in a distributed manner, saving content from remote locations securely through the Internet or your company intranet.

Automated Data Capture

XenDocs supports a wide-range of automated data capture tools to extract meaningful information from scanned documents and forms. Using Optical Character Recognition (OCR) and related technologies, XenDocs can "read" machine typed and hand printed data from nearly any scanned image.

Unstructured Documents

Unstructured documents such as legal contracts, memos and mailroom documents contain data that cannot be mapped to a database, so users generally need access to the entire document in

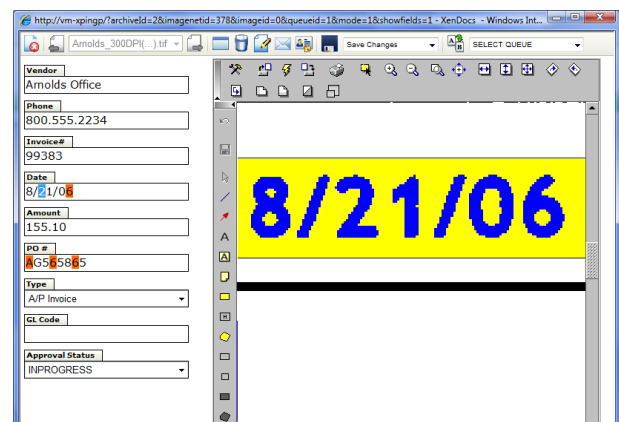
order to access any of the information it contains. With unstructured documents, XenDocs performs "full-text" OCR to capture all of the information and populate a full-text index. This enable users to perform keyword searches on the entire document.

Forms

Specific data elements can be automatically extracted from structured and semi-structured forms. The extracted data is then routed for validation and correction by a human operator. ***This process typically eliminates 50-90% of the time required, versus manual data entry.***

With **structured forms** like timecards, surveys, CRFs, applications and enrollments, XenDocs extracts data based on templates with zonal OCR fields. These templates capture hand printed data, marked bubbles, checkboxes and barcodes.

Semi-structured forms like A/P invoices and customer purchase orders have consistent data schemas, but varying layouts on each document. XenDocs uses flexible OCR templates to capture the machine printed data that "floats" on each document but can be located relative to predefined keywords.



Web-based OCR Correction

The web-based OCR correction module (patent pending) within XenDocs blends the responsiveness of desktop software with the accessibility of the web. This ergonomic, thin-client interface automatically selects only the "low confidence" characters for review, providing fast and accurate processing by your staff.

It is optimized for organizations with distributed teams of users who must validate OCR results or review documents.