BECKER&POLIAKOFF

contentCrawler from DocsCorp replaced a costly, cumbersome four-step OCR'ing workflow process at Becker & Poliakoff, streamlining it into an efficient single-step process. OCR processing – which used to be expensive and time-consuming for the firm – is now accomplished within a centralized framework, saving time and allowing users to stay in control of their workspace.

Executive Summary

Client Challenge

Transform the existing lengthy and expensive OCR process into a single-step workflow

Improve efficiency by automating the OCR process

contentCrawler as a Solution

Reduced time spent OCR'ing documents

Delivered a backend process for automatically OCR'ing files

Found non-indexable files in their DMS and processed them with OCR

Business Value

Supports faster, more efficient processes that reduces the amount of time spent OCR'ing documents

Allows users to stay in control of what goes into their workspace

Runs as a fully-automated OCR processing service 24/7 with no staff intervention required

Is a cost-effective solution that replaces both hardware, in the form of scanners, and software

Finds and OCRs legacy documents while also processing newly profiled documents

DocsCorp

About Becker & Poliakoff

Becker & Poliakoff, with headquarters in Fort Lauderdale, Fla., is a multi-practice commercial law firm with attorneys, lobbyists and other professionals at offices across the United States.

Challenge: Make every document text-searchable

Staff at Becker & Poliakoff were using a four-step process to convert paper documents to text-searchable PDFs. They would scan the documents into their system, open the documents in pdfDocs (a PDF creation and editing application from DocsCorp), OCR the documents by adding a text layer, and profile the new text-searchable PDF into their DMS.

Though this workflow helped meet regulatory requirements around storing files as PDFs, the whole process was extremely time-consuming and error-prone. If staff didn't follow all the steps in the process or skipped the process entirely, the documents could easily be profiled into the DMS as non-searchable documents.

Solution: An automated, end-to-end OCR framework

Becker & Poliakoff IT staff knew they needed an automated end-to-end system that could handle the conversion of documents to text-searchable PDFs more efficiently. They reached out to DocsCorp to deploy contentCrawler.

contentCrawler is an integrated analysis, processing and reporting tool that assesses image documents in content repositories to apply OCR conversion. contentCrawler's automated end-to-end process converts image documents to text-searchable PDFs, which are saved back into the content repository - ready to be found by searching. This approach reduces non-compliance risks and increases efficiency.

Becker & Poliakoff decided to run contentCrawler in Active Monitoring and Backlog modes simultaneously. Active Monitoring performed OCR processing on all newly profiled documents in the DMS. Backlog processing, on the other hand, converted all the legacy documents in the system to text-searchable PDFs. This process took longer as there were hundreds of thousands of documents to be assessed and converted to PDF. contentCrawler can work through 17,000 pages per day on a 4-core machine. Firms can add additional cores to speed up the process.

Summary

Having non-searchable documents in its DMS was an unacceptable and unnecessary risk for Becker & Poliakoff. To resolve the issue, the firm deployed contentCrawler – an automated, end-to-end OCR framework from DocsCorp. Now, all documents saved into their DMS are fully text searchable and compliant with industry requirements.