

New Yorker Behörde verbessert OCR-Effizienz um 85% mit der PDF Optimization Suite für OpenText Captiva®



Geschäftliche Herausforderung

Das New York State Department of Environmental Conservation (NYSDEC) musste bislang alle relevanten Dokumente mittels einer Suchanfrage finden, mit der Text innerhalb von Dateien, über alle Dateien hinweg gesucht wurde. Zu diesen Dokumenten gehören komplexe PDF-Dateien – elektronische, Bild- und Hybriddokumente. Die generische OCR-Technologie erforderte, dass die PDF-Dateien in Bilder umgewandelt und anschließend mit OCR verarbeitet wurden. Dabei gingen wichtige Informationen wie Lesezeichen, Inhaltsverzeichnisse oder Hyperlinks verloren. Die staatliche Behörde musste sicherstellen, dass alle Texte innerhalb eines Dokuments indiziert wurden und die wichtigsten PDF-Eigenschaften erhalten bleiben, um die Vollständigkeit ihrer wichtigen Dokumente zu bewahren.

Lösung

Um alle Dokumente zu verschlagworten und die in den Dokumenten der Behörde enthaltenen Informationen zu erhalten, wechselte das NYSDEC zu Foxits PDF Optimization Suite für OpenText Captiva®. Die Suite wurde verwendet, um alle Dokumente der Behörde in ihrem System, (OpenText Documentum®) im Volltext, über alle Inhalte des NYSDEC ohne Informationsverlust, durchsuchbar zu machen. Darüber hinaus wurden die mit OCR verarbeiteten Dokumente vollständig indiziert, sobald sie wieder in ihre Documentum-Umgebung eingefügt wurden.

Um **85%** verbesserte OCR Verarbeitungsgeschwindigkeit



Aufgrund der Vermeidung der unnötigen Umwandlung in Bilder konnte mit der PDF Optimization Suite für OpenText Captiva® die OCR-Effizienz in der Behörde deutlich gesteigert werden.



Ergebnisse

Die PDF Optimization Suite für OpenText Captiva® von Foxit war in der Lage, das Problem des Informationsverlustes durch den zuvor eingesetzten, generischen OCR-Prozess zu lösen und gleichzeitig vollständig indizierte Dateien zu erstellen. Als Experte für PDF-Technologie hat Foxit eine Lösung entwickelt, die es überflüssig macht, PDF-Dateien vor der Verarbeitung in Bilder umzuwandeln. So wird der Verlust wichtiger Eigenschaften, wie Lesezeichen, Metadaten, PDF/A, Hyperlinks etc. verhindert. Ohne die unnötige Bildumwandlung konnte die PDF Optimization Suite für OpenText Captiva® zudem die Verarbeitungszeit auf nur 15% der Zeit reduzieren, die der bisher verwendete generische OCR-Prozess benötigte. Diese Lösung half dem NYSDEC dabei, alle wichtigen Informationen zu erhalten und gleichzeitig die Verarbeitungsgeschwindigkeit deutlich zu verbessern. Zudem wurde die PDF Optimization Suite für OpenText Captiva® schnell und einfach in den bestehenden Workflow integriert.

Vorteile

Zielvorgabe	Erreichte Vorteile
<ul style="list-style-type: none">▶ Eliminieren von Informationsverlusten, die durch den vorherigen, generischen OCR-Prozess verursacht wurden	<ul style="list-style-type: none">✓ Die Foxit PDF Optimization Suite für OpenText Captiva® war in der Lage, die Dokumente der Behörde unter Beibehaltung wichtiger Eigenschaften wie Lesezeichen, Metadaten, PDF/A, Hyperlinks usw. mit OCR zu verarbeiten.
<ul style="list-style-type: none">▶ OCR-Verarbeitung aller Dokumente der Behörde, einschließlich komplexer PDF-Dateien	<ul style="list-style-type: none">✓ Die PDF Optimization Suite für OpenText Captiva® machte das NYSDEC-Archiv vollständig und ohne Informationsverluste durchsuchbar.
<ul style="list-style-type: none">▶ Vermeiden der unnötigen Umwandlung von PDF-Dateien in Bilder vor der Verarbeitung	<ul style="list-style-type: none">✓ Durch die Vermeidung der Bildumwandlung konnte die PDF Optimization Suite die Zeit für die OCR-Verarbeitung auf nur 15% der Zeit reduzieren, die der bisher eingesetzte generische OCR-Prozess benötigte.
<ul style="list-style-type: none">▶ Nahtlose Integration in die vorhandene Workflow-Umgebung	<ul style="list-style-type: none">✓ Die PDF Optimization Suite für OpenText Captiva® lässt sich nahtlos sowohl in Documentum- als auch in Captiva-Umgebungen integrieren.