

iSquare nutzt Technologie von callas für PDF/A-konform archivierte Tageszeitungen bei der Deutschen Nationalbibliothek

(Berlin) Der Softwareanbieter iSquare aus Berlin hat sich im Rahmen eines Projektes für die Deutsche Nationalbibliothek zur Nutzung der pdfToolbox von callas software entschieden. Das Unternehmen stellt seit Juni 2010 täglich die sogenannten e-Paper Ausgaben von 300 deutschen Tageszeitungen in PDF/A bereit, die die Nationalbibliothek zur langfristigen Archivierung und Sicherung für die Nachwelt benötigt. Für die korrekte Erzeugung der Archivdateien sorgt unter anderem die pdfToolbox. Sie konvertiert im Rahmen eines Batch-Prozesses die von iSquare per Spiderprogramm über Nacht gesammelten Ausgaben in qualitativ hochwertige PDF/A-Dateien.

Zum Auftrag der Deutschen Nationalbibliothek gehört es unter anderem sicherzustellen, dass die Inhalte der deutschen Tageszeitungen langfristig archiviert werden. Daher wurde Ende 2009 in einem Ausschreibungsverfahren ein Dienstleister für die „Sammlung, Konvertierung und Bereitstellung elektronischer Ausgaben von Tageszeitungen“ gesucht. Den Zuschlag erhielt der Softwareanbieter iSquare. Da die Ausschreibung die Konvertierung aller elektronischen Zeitungen in das Langzeitarchivformat PDF/A vorschrieb, integrierte iSquare nach umfangreicher Marktsondierung unter anderem die pdfToolbox von callas in die Gesamtlösung.

„Die Technologie von callas software hat uns insbesondere durch die hohe Qualität der Ergebnisse überzeugt“, erklärt Michael Kapst, Geschäftsführer von iSquare. „Damit ergänzt die pdfToolbox optimal unsere Lösung für die Deutsche Nationalbibliothek.“

Komplexer Prozess erfordert zahlreiche Einzelschritte

Um die elektronischen Ausgaben der Deutschen Nationalbibliothek kontinuierlich zur Verfügung zu stellen, sind umfassende Arbeiten notwendig. Zunächst holt der iSquare ePaper Manager zwischen 8.00 und 10.00 Uhr die elektronischen Ausgaben über ein eigenes Spiderprogramm auf den

Internetseiten der Tageszeitungen ab. „Die Verlage bieten die ePaper auf verschiedenen Wegen zum Download an“, erläutert Kapst. „Bei den Verlagen muss man oft die Seiten auswählen, die in die PDF-Dateien integriert werden sollen. Dazu kann man entweder direkt die jeweiligen Seiten oder aber Rubriken, wie Politik, Sport usw. markieren. Zusätzlich muss bei vielen Verlagen vorher die Wahl auf die Regionalausgabe getroffen werden.“

Die Zeitungsverlage sind verpflichtet, ihre elektronischen Zeitungen in einem konvertierbaren Format bereitzustellen. Ist dies nicht der Fall, wird dies protokolliert und gesondert reklamiert. Zu jeder erfassten Ausgabe wird außerdem automatisiert ein XML-Metadatenatz erstellt. Schließlich werden die Ausgaben insbesondere auf Vollständigkeit und Konsistenz geprüft und dann in einer Datenbasis abgelegt, auf die via Weboberfläche zugegriffen werden kann. Hier stehen sie zur weiteren Bearbeitung und eventuellen Fehlerbearbeitung zur Verfügung. Anschließend findet die Konvertierung nach PDF/A statt. Hierfür hat iSquare die pdfToolbox in ihre Lösung integriert, die für die vollumfängliche Einhaltung sämtlicher Spezifikationen von PDF/A sorgt und Fehler automatisiert korrigiert. Die pdfToolbox analysiert die PDF-Dokumente und greift bei Bedarf auch in eingebettete Objekte wie zum Beispiel Schriften, Metadaten oder Bilder oder Kompressionsalgorithmen ein. So werden alle versteckten Probleme erkannt und bei der Konvertierung behoben.

Die im Schnitt täglich 8.400 Seiten stehen in der Regel bis mittags der Deutschen Nationalbibliothek samt der Metadaten über eine OAI (Open Archives Initiative)-konforme Schnittstelle zur Abholung bereit.

„Für das Institut spielt der Zeitfaktor allerdings keine große Rolle, wichtiger sind Vollständigkeit und Korrektheit der Dateien. Und dazu trägt auch unsere pdfToolbox maßgeblich bei“, erklärt Olaf Drümmer, Geschäftsführer von callas software abschließend.

Über callas software

callas software bietet einfache Möglichkeiten, um komplexe Herausforderungen im PDF-Umfeld zu bewältigen. Als Innovator von Verfahren entwickelt und vermarktet callas software PDF-Technologien für das Publishing, die Produktionsstufe, den Dokumentenaustausch und die Dokumentenarchivierung.

callas software unterstützt Agenturen, Verlagen und Druckereien darin, ihre Aufgabenstellungen zu lösen, indem es Software zum Prüfen, Korrigieren und Wiederverwenden von PDF-Dateien für die Produktion von Druckvorlagen und das elektronische Publizieren bereitstellt.

Unternehmen und Behörden aus aller Welt vertrauen den zukunftssicheren, vollständig PDF/A-konformen Archivierungslösungen von callas software.

Darüber hinaus ist die Technologie von callas software auch als Programmierbibliothek (SDK) für Entwickler erhältlich, die PDFs optimieren, validieren und korrigieren müssen. Software-Anbieter wie Adobe®, Quark®, Xerox® und viele andere haben verstanden, welche Qualität und Flexibilität die callas-Werkzeuge bieten, und haben sie in ihre eigenen Lösungen integriert.

callas software unterstützt aktiv internationale Standards und wirkt in ISO, CIP4, der European Color Initiative und der Ghent PDF Workgroup mit. Darüber hinaus ist callas software Gründungsmitglied des PDF/A Competence Center. Der Firmensitz des Unternehmens befindet sich in Berlin. Weitere Informationen finden Sie im Internet unter www.callassoftware.com.

Ihre Redaktionskontakte

callas software GmbH

Dietrich von Seggern
Schönhauser Allee 6/7
D-10119 Berlin
Tel +49 30 443 90 31-0
Fax +49 30 441 64 02
d.seggern@callassoftware.com
www.callassoftware.com

good news! GmbH

Nicole Körber
Kolberger Str. 36
D-23617 Stockelsdorf
Telefon: +49 451 88199 - 12
Telefax: +49 451 88199 - 29
nicole@goodnews.de
www.goodnews.de