



ARCHIVIERUNG MIT AUTOPILOTEN

WIE INTELLIGENTE DOKUMENTE FÜR EINE AUTONOME ARCHIVIERUNG SORGEN

Die Automobilindustrie hat es vorge-macht: Autonomes Fahren bedeutet mehr Freiheit für den Fahrer. Analog dazu, sorgt das autonome Archivieren für freie Kapazitäten und Kosteneinsparungen in der IT. Wer eine intelligente und „aufge-räumte“ Archivierung mittels Autopiloten anstrebt, kann jetzt loslegen.

Vom autonomen Fahren, also der fahrer-losen Fortbewegung mit einem Fahr-zeug, hat heute jeder eine Vorstellung. Schon bei der Erfindung des Automobils wurde durch die Namensgebung (autos = selbst, persönlich, eigen / mobilis = beweglich) das „Selbstbewegliche“ be-tont, auch wenn die Autonomie des Fahr-ers durchaus überschaubar war. Rück-blick: Hinsichtlich der Autonomie in der Fortbewegung gab es zunächst die mit-telbare Abhängigkeit vom Pferd gefolgt von der absoluten Abhängigkeit vom Menschen. Doch in Zeiten von Tesla und Co gibt es sie nun tatsächlich, die Steue-

rung eines Fahrzeugs ohne das Zutun des Menschen. In Bezug auf die Archi-vierung lassen sich frappierende Analog-ien erkennen: absolute Abhängigkeit vom Menschen (Organisation von Pa-pierordnern), vielfache Abhängigkeit vom Menschen (Organisation von digi-talisierten Ablagen) und nun befindet sich der Mensch auf dem Weg zur auto-nomen Archivierung.

Das intelligente Dokument

Es sind nicht die Straße oder der Zielort, die das Auto zum autonomen Fahren be-fähigen, sondern das Auto selbst. Wäre es beispielsweise die Straße, würde es sich wie beim Schienenverkehr verhalten und wäre weit weg von jeglicher freien Bewegung und Flexibilität. Genau dies lässt sich auch auf das Thema Dokumen-tenarchivierung übertragen: Weder die Software-Anwendung, in der das Doku-ment erstellt und bearbeitet wird, noch der Speicher, in dem das Dokument ar-

chiviert wird, können die Archivierung steuern. Eine dritte Ebene, so wie es häu-fig praktiziert wird – ein DMS (Dokumen-ten Management System) oder ECM (En-terprise Content Management System) – wäre ebenso eine übergeordnete Intelli-genz. Diese bringt erstens weiteren „Systemballast“ mit sich und raubt zwei-ten ebenfalls Flexibilität. Der Grund: Sys-teme gehen, aber die Dokumente blei-ben. Der Austausch einer Instanz wird häufig zum Großprojekt für die IT.

Das intelligente Dokument ist die neue Lösung für diesen Flickenteppich der Sys-teme. Dokumente tragen diverse Informa-tionen in sich, wie beispielsweise den Content selbst, aber auch Meta- und Kon-textbezogene Daten. Die Metadaten et-wa stammen aus den führenden Anwen-dungen, werden dem Dokument sozusagen „aktiv“ mitgegeben. Beim Abschalt-en einer Anwendung, gehen diese Daten im Falle des intelligenten Dokuments nicht

mehr verloren. Vielmehr machen sie das Dokument zu einer von der Anwendung unabhängigen, autonomen Instanz, die jederzeit in anderen Anwendungen angezeigt und bearbeitet werden kann.

Autopilot/Autopilotin

Wer macht das Dokument intelligent? Ein System, das die Daten des Dokuments an geeigneter Stelle erfasst, auslesen und weitergeben kann, um die „freien“ Dokumente sozusagen in die richtigen Bahnen zu lenken. Die Managerin oder Autopilotin, die diese Anforderungen realisiert, nennt KGS tia – the intelligent archive. Schnittstellentechnisch setzt das intelligente Archiv dabei auf den Content Management Interoperability Services-Standard (CMIS). Anders als etwa die SAP Archive-Link-Schnittstelle ermöglicht er die Verwaltung von Metadaten; eine weitere dafür geeignete Schnittstelle ist OData, -verbreitet vor allem im SAP-Umfeld, aber auch bei Anwendungen wie MS SharePoint.

tia schafft über das intelligente Dokument in der Folge einen Single Point of Truth (SPoT) für Dokumente jeder Art. Ganz gleich, in welcher Anwendung gearbeitet wird – auf die Dokumente kann von überall zugegriffen werden.



„DAS ARCHIV DER ZUKUNFT MUSS DOKUMENTE MIT INTELLIGENZ AUFLADEN, MUSTER SELBST ERKENNEN UND VORAUSSCHAUEND OPTIMIERUNGSVORSCHLÄGE BRINGEN.

Benjamin Schröder, Head of R&D,
KGS Software, www.kgs-software.com

Autonome Archivierung

Das Archiv der Zukunft muss Muster selbst erkennen und vorausschauend Optimierungsvorschläge bringen. Übervolle Archive sollten reorganisiert, Vorschläge zum Löschen gemacht, teure Speicherplätze entlastet und günstiger Speicher umfänglich genutzt werden. Dies alles jedoch nicht wahllos, sondern diversen Informationen des intelligenten Dokuments folgend. Ein rechtlich relevantes Dokument muss demnach anders „gelagert“ werden als eine Alltagsnotiz. Jedes Kriterium, jede Metainformation, jede Frist – das alles kann unmöglich jedem einzelnen Dokument „aktiv“ via Metadaten mitgegeben werden. Auf dem Weg zu einer vollständigen Automatisierung der Prozesse aber benötigt das intelligente Archiv genau diese Informationen, die es lernen lassen und Handlungen auslösen. Die spannendsten Informationen sind dabei häufig diejenigen, die das intelligente Dokument selbst sozusagen „inaktiv“ erlebt und die es inhaltlich ausmachen. Ein lernfähiges System kann diese Daten aufnehmen und auf deren Grundlage vielleicht sogar besser - sehr wahrscheinlich jedoch verlässlicher und unfehlbarer – handeln, als der Mensch.

Diese Fähigkeiten lassen sich über Statistics realisieren. Zunächst beschreibt Statistics nur die Disziplin, die die Erfassung, Organisation, Analyse, Interpretation und Darstellung von Daten betrifft. Es ist jedoch die Basis für regelbasierte Operationen und ebnet den Weg zu einem selbstlernenden System (Machine Learning) und den vorausschauenden Services.

Ordnung muss sein

Die Kontext- und Contentdaten sind das eigentliche Gold, nach dem ein intelligentes Archiv schürft. Ein Beispiel sind Nutzungsdaten, von wem, wann und wie oft wurde das Dokument zurück aus dem Archiv auf den Bildschirm geholt. Dokumente, die beispielsweise fünf Jahre nicht geöffnet wurden, können zum Löschen vorgeschlagen oder automatisch und transparent auf einen günstigeren Speicher umgezogen werden.



„DAS INTELLIGENTE DOKUMENT SCHAFFT UNABHÄNGIGKEIT – DENN SYSTEME GEHEN, ABER DIE DOKUMENTE BLEIBEN.

Johanna Zinn, Head of Marketing,
KGS Software, www.kgs-software.com

Technisch funktioniert dies über sogenannte Data-Lakes. Wie der Begriff schon andeutet, handelt es sich wahrlich um einen See von Daten, also einen sehr großen unstrukturierten Datenspeicher. Mit Zugriff auf diese Lakes können Analysen und diverse Operationen durchgeführt werden. Dabei werden Daten erst strukturiert oder gegebenenfalls umformatiert, wenn sie konkret benötigt werden. Neben text- oder zahlenbasierten Daten können Data-Lakes auch Bilder, Videos oder andere Datenformate aufnehmen. Ein Content bezogenes Beispiel aus dem Personalbereich und dem Datenschutz könnte demnach sein: Ex-Mitarbeiter X verlangt eine Löschung aller Fotos mit seinem Konterfei von Betriebsfeiern und offizieller Nutzung. Nun muss entsprechend auf Bilder- und Videodaten zurückgegriffen werden aus denen entsprechende Inhalte extrahiert werden.

Viele große und mittelständische Unternehmen haben bereits Data-Lakes beziehungsweise ganze Datensysteme für ihre Business Intelligenz Anforderungen aus anderen Bereichen im Einsatz. Die autonome Archivierung mit tia zeichnet sich dadurch aus, dass sich die Archiv-Management-Fähigkeiten auch flexibel in diese bestehenden Systeme von zum Beispiel Google, AWS und SAP einfügen lassen.

Benjamin Schröder, Johanna Zinn