

# Die Eingangspost-Digitalisierung hält Einzug

In immer mehr Unternehmen wird über den Start in die elektronische Eingangspost diskutiert: ein Leitfaden.

*Papier bremst das Geschäft. Das ist nicht nur für Workflow-Experten ein Allgemeinplatz, sondern entspricht auch der Erfahrung an jedem Sachbearbeiter-Platz. Es ist aber gar nicht so einfach, daraus die richtigen Konsequenzen zu ziehen. eMails und Telefaxe, digitale EDI-Informationen und immer wieder jede Menge Papier – wie soll man diese unterschiedlichen Medien auf einen Nenner bringen? In immer mehr Unternehmen und Organisationen ist der Digital Mailroom die Lösung: Die klassische Poststelle wird zum Eingangsportale des digitalen Dokumentenmanagements umgebaut. Dabei muss man aber auf eine Menge Dinge Acht geben.*



POSTMASTER-Magazin (2)

**DMS: Papier erkannt – Papier verbannt**

Wo in Firmen und Behörden neue Techniken angewandt werden, da will man Abläufe verbessern und Kosten sparen. Als in den Poststellen die ersten Scanner auftauchten, gab es die klare Erwartung, dass damit die Arbeit im Eingangspost- und Hausservice leichter von der Hand geht, dass die Eingangspost mit weniger Arbeit trotzdem schneller in die Organisation kommt. Das hat fast nirgendwo geklappt. Im Gegenteil. Wenn Dokumente zentral so aufgearbeitet werden, dass sie für die Arbeit in der Organisation optimal sind (und nach der Archivierung auch noch problemlos wiedergefunden werden können), dann gibt es für die elektronische Eingangs-

postbearbeitung eine Menge neuer Aufgaben.

## Scannen – lesen – klassifizieren

Papierdokumente muss man scannen, damit sie in elektronischer Form vorliegen. Das geht ziemlich schnell, aber leider fängt danach die Arbeit erst richtig an. Der Inhalt muss analysiert und zugeordnet, der Zusammenhang zum Kunden oder Bürger, zum Projekt oder Schadensfall, zum Auftrag oder zur Bestellung muss hergestellt werden. Eine Technik muss her, die in der Lage ist, den Inhalt eines Dokuments zu ergründen – gleichgültig ob es als eMail, als Brief oder Fax eingetroffen ist.

Heute kann man mit OCR (Optical Character Recognition) beim Scanvorgang schon für die elektronische Zeichenerkennung sorgen. Wenn die Buchstaben entschlüsselt sind, müssen Text, Layout und Struktur des Dokuments analysiert werden, um herauszufinden, wie damit weiter umzugehen ist.

Der Traum aller Klassifikations-Profis ist die erfolgreiche „Dunkelverarbeitung“: Ohne manuellen Eingriff werden hier Dokumente als Rechnungen erkannt, Kundennummer, Bestellmenge und Rechnungsbetrag mit den digital abgespeicherten Werten von Bestellung und Wareneingang verglichen und das Ganze dann mit 99,99-prozentiger Sicherheit sofort an die Buchhaltung und die Abteilung Materialwirtschaft digital weitergeleitet. Auch wenn der Digital Mailroom gut funktioniert, klappt das heute noch nicht immer – aber immer öfter.

## Die Aufgabe bestimmt die Lösung

Besonders gut funktioniert das Erfassen und Zuordnen der Eingangsdokumente, wenn es sich um Formulare handelt, die die Organisation selbst entworfen und den Partnern draußen zur Nutzung übergeben hat. Dann kann die formularbasierte Software die verschiedenen Areale des Dokuments ansteuern und die dort eingetragenen Informationen zuordnen. Wenn es hier beim Ausfüllen keine Fehler gegeben hat, sind gute Leseergebnisse zu erwarten.

Bei vielen Dokumentenarten gibt es aber keine feste Struktur. Dann muss die Lesesoftware sich komplexe Informationen wie Bestellnummern, Adressdaten oder Rechnungspositionen aus dem erfassten Text nach definierten Regeln herausuchen. Die Freiform-Erkennung braucht natürlich mehr Zeit und Rechnerkapazität, da sie ohne feste Struktur des Dokuments auskommen muss. Stattdessen

versucht sie, Regeln anzuwenden. Zum Beispiel:

- Eine achtstellige Zahl, in der die ersten drei, dann nochmals drei und dann zwei Ziffern zusammengefasst sind, ist mit einer Wahrscheinlichkeit von X eine Bankleitzahl.
- Die Wahrscheinlichkeit erhöht sich um ein Vielfaches, wenn vor einer so gruppierten Zahl die Buchstabengruppe BLZ steht.
- Zahlengruppen nach Texten wie Rg. Nr., Rechnung Nummer oder Aktenzeichen sind mit einer Wahrscheinlichkeit von X Rechnungsnummern.
- Bei einer Zahlenfolge, die aus zwei Ziffern, einem Bindestrich, danach vier Ziffern, danach noch einem Bindestrich und zum Schluss weiteren zwei Ziffern besteht, handelt es sich wahrscheinlich um die Schadennummer, wie unsere Versicherung sie vergibt.

Es gibt aber auch Dokumente, die sind keine Formulare, und sie betreffen auch nicht standardisierte Geschäftsvorfälle wie Angebote, Aufträge oder Rechnungen. Hier hat es überhaupt keinen Zweck, dass sich die Klassifikations-Software mit der Textanalyse abarbeitet.

### Die Klassifizierung von Dokumentenarten

ist ein sinnvoller Arbeitsgang, mit dem man viel unnütze Rechnerzeit und noch mehr Irrtümer verhindern kann. Wie man die eingehenden Dokumente und Vorgänge sinnvoll sortiert, zeigt wahrscheinlich die Erfahrung. Eine Dreiteilung in Formulare, Standard-Geschäftsvorgänge und sonstige Briefe ist aber wohl auf jeden Fall sinnvoll. Man könnte annehmen, dass die bereits in digitaler Form vorliegenden eMails einfacher zu klassifizieren sind als gescanntes Papier. Das stimmt aber oft nicht. eMails sind häufig noch wesentlich formloser und schwieriger zu erschließen als relativ ordentlich aufgebaute Geschäftsbriefe. Ein ausgesprochener OCR- und Capturing-Killer sind die Riesen-eMails, die nach mehrfacher, beiderseitiger Beantwortung entstanden sind. Kunden- und Auftragsnummern sowie Adressen sind hier entweder gar

nicht vorhanden oder in einem Wust von Texten versteckt.

### Verbesserung der Erkennungsrate

Jeder Automatisierungsschritt ist besser als die „Klassifikation zu Fuß“ am Bildschirm. Dafür ist es nützlich, dass System laufend zu trainieren:

- Vergleiche die gelesene Bestellnummer mit den Bestellnummern aus unserem Materialwirtschafts-System.
- Suche eine Entsprechung zu der im Dokument gefundenen Kundennummern in der Partnerdatenbank; stim-

men Name, Postleitzahl oder Ort in beiden überein?

- Beim Lieferanten X steht die Rechnungsnummer immer unter dem Text: „Bei Überweisungen immer angeben.“
- Eingangsdokumente, in dem der Begriff „Reklamation“ gefunden wird, müssen immer manuell weiterverarbeitet werden.

Außer der Vorsortierung nach Dokumententypen kann man auch durch bestimmte Wünsche oder Vorgaben an Partner ein paar Verbesserungen erreichen. Es hilft zum Beispiel, wenn Lieferanten wissen, dass sie pro Bestel-

## Erkennungs-Software

Anbieter	Formular-Verarbeitung	Freiform-Verarbeitung
Autonomy Cardiff	TeleForm	
Abbyy	FormReader	FlexiCapture
AnyDoc	OCR for AnyDoc	AnyDoc Invoice
BancTec	eFirst Capture	eFirst Capture
Beta Systems	FormsRec	FrontCollect Invoice / FrontCollect Mailroom
B&L OCR-Systeme	FormStar, AdressStar	DokuStar
Open Text	Capture Recognition Engine (RecoStar)	Capture Document Reader (DokuStar)
Comline		helic Recognition
d.velop		d.classify
Docutec	X4D (Xtract for documents)	X4D (Xtract for documents)
Easy Software		Easy Capture mit Easy IQ Classify
Elo Digital Office	DocXtractor II	DocXtractor II
EMC Captiva	FormWare	Captiva Dispatcher / Captiva InputAccel
Future Software	JustScan Enterprise	JustScan Next Generation
I.R.I.S Group	IrisCapture Pro for Forms	IrisCapture Pro for Invoices
Insiders Technologies	smartFix	smartFix
Isis Papyrus	Papyrus FixForm	Papyrus FreeForm / Papyrus Classify
Itesoft	FreeMind for Business	FreeMind Enterprise
K7 it Solutions	Wando	Wando invoice / Wando info center
Kofax	Transformation Modules	Transformation Modules
Paradatec	Prosar-Aida	Prosar-Aida
ReadSoft	Documents for Forms	Documents for Invoices / Documents for Mailrooms
Saperion		Inbound Suite
Scalaris	Application Forms	Invoice, Letter Post, Files & Correspondence
Seeburger		4invoice
SER Solutions	Doxis Inbound Master	Doxis Inbound Master / Doxis Invoice Master
Top Image Systems	eFlow Integra	eFlow Freedom / eFlow Smart
TCG Informatik	FreeFormStar	FreeFormStar

Quelle: Zöller & Partner



**Kunterbuntes Beleggut ist für moderne Produktionsscanner kein Problem.**

lung eine einzelne Rechnung ausstellen sollen, und wenn auf jedem Lieferschein eine Auftragsnummer steht. Wenn – wie zum Beispiel bei Versicherungen – stark mit vorgefertigten Formularen gearbeitet wird, dann hat man das Optimierungspotenzial selbst in der Hand: Wo direkt maschinenlesbare Codes und Kennungen verwendet werden können, hat der Digital Mailroom erheblich leichteres Spiel. Wo beim Ausfüllen Kästchenreihen oder Ausfüllkämme verwendet werden, da sinkt die Zahl der Formularnutzer, die fließende Schreibrschrift verwenden. Besonders gut ist natürlich die Bereitstellung elektronischer Formulare zum Ausfüllen und Ausdrucken. Dokumente bestehen oft aus mehr als einem A4-Blatt. Auch eine leistungsfähigere Capture-Software kann oft nicht ohne weiteres erkennen, wo ein Dokument anfängt und wo es aufhört. Wo die Software-Regeln über die Erkennung von Dokumentengrenzen nicht richtig greifen, da helfen Trennblätter.

### „Erkennungsrate 99 Prozent“

Der Maßstab für die Leistungsfähigkeit einer Klassifikations-Software ist natürlich deren Fähigkeit, Dokumente zu erkennen und zuzuordnen. Wenn hier mit hohen und höchsten Prozentsätzen argumentiert wird, sollte man genau hinschauen:

- Geht es um die Zahl der erkannten Belege? Dann bedeutet 99 Prozent tatsächlich, dass von 100 Eingangs-

dokumenten nur eines nicht korrekt verarbeitet wurde. Das ist ein sehr guter Wert.

- Wenn – speziell bei der Formularverarbeitung – ein Prozent von tausend Feldern nicht identifiziert wurden, die es auf diesen hundert Dokumenten zu erkennen gab, dann sind im ungünstigen Fall zehn von hundert Dokumenten (wegen Identifikationsproblemen in je einem Feld) nicht korrekt verarbeitet worden. Das ist schon weniger gut.
- Wenn schließlich von 10000 Zeichen, die auf den hundert Dokumenten zu identifizieren waren, ein Prozent nicht erkannt wurden, dann konnte im statistisch ungünstigsten Fall kein einziges Dokument fehlerfrei identifiziert und klassifiziert werden. Das ist offensichtlich ein ganz und gar schlechtes Ergebnis.

Hinter hohen Erkennungsraten können also ganz unterschiedliche Phänomene stecken oder versteckt werden. Softwarelösungen für die Formular- und Freiform-Erkennung gibt es in großer Zahl. Insgesamt sind gewiss zwei Dutzend Anbieter auf dem Markt. Die wichtigsten Namen finden Sie im Infokasten auf der vorigen Seite.

### Erst der Test, dann die Entscheidung

Wer für das zukünftige Capturing und die Dokumentenklassifikation verantwortlich ist, der sollte sich als Testwerkzeug einen für das eigene Unternehmen typischen Dokumentenstapel zurechtlegen, den man als Testfall mit den Systemen der verschiedenen Anbieter laufen lassen kann. Nur wenn der Fall bei den Probeläufen immer gleich bleibt, ist ein echter Vergleich möglich. Profis arbeiten sogar „doppelstöckig“: Zunächst wird ein Dokumentenstapel an den Hersteller übergeben, damit er eine Chance hat, das System auf die Struktur des Teststapels hin zu optimieren. Interessant wird dann ein zweiter Testlauf mit einem ebenfalls repräsentativen Dokumentenstapel. Hier kann man bereits prüfen, wie die Optimierungen aus dem ersten Testlauf sich auf die folgende Verarbeitung auswirken.

### „Alles ganz anders bei uns“

Weil die Probleme bei der Eingangspost-Digitalisierung und der anschließenden elektronischen Zustellung in jeder Organisation anders aussehen, schieben viele Firmen diese nützlichen Modernisierungsmaßnahmen vor sich her: „Viel zu kompliziert – wie sollen wir so etwas jemals ans Laufen bringen? Das ist doch alles ganz anders bei uns!“ Tatsache ist, dass die Verwandlung unterschiedlichster Informationen in ein durchgängiges Medium für die gemeinsamen Arbeitsabläufe eine gewaltige Erleichterung ist, von der alle Instanzen profitieren. Ganz hohe Prozentsätze der bisherigen Sucherei im Papier können entfallen. Auf Dokumente kann gemeinsam und gleichzeitig von verschiedenen Abteilungen und Instanzen zugegriffen werden, und schließlich wird dann auch die Archivierung der jederzeit vollständig vorhandenen elektronischen Vorgangsakte ein Kinderspiel. Als Poststellenleiter sollte man insbesondere einen wichtigen Gesichtspunkt nicht übersehen: Wer Eingangspost digitalisiert, dessen Funktion geht ganz eindeutig über das Sortieren und Austragen von Eingangspost und Akten hinaus. Wer klassifizierte Informationen und Dokumente ins Dokumentenmanagementsystem hineingibt, der stößt Workflows an und ist wichtige Ausgangsinstanz für die Durchführung betrieblicher Abläufe. Wer im Betrieb an der Zukunft der Eingangspostverwaltung arbeitet, muss diese wichtige Chance ergreifen. Nur was die Arbeitseinsparung in der Poststelle angeht – dazu sollte man keinerlei Versprechen machen.

### Die Hardware-Schnittstelle

Der Scanner schafft den Medienbruch vom Brief zur Datei. Ein rundes Dutzend Hersteller liefern leistungsfähige Produktionsscanner, die in der Lage sind, den täglichen Posteingang zentral zu bewältigen. Die Geräte sind in der Lage, Mischpost mit unterschiedlichsten Anlagen automatisch einzuziehen und zu verarbeiten. Auf den folgenden Seiten stellen einige dieser Hersteller ihre Lösungen für den Digital Mailroom vor.

*bdw*