

Capturing in der Poststelle

Eine Übersicht über Dokumenten-Erfassungssoftware – Teil 1

Immer mehr Firmen denken über die Anschaffung eines komplexen Dokumentenmanagement-Systems nach. Mit allen Funktionen vom Posteingang über den Workflow bis ins Archiv inklusive eMail-Einbindung ist das allemal ein Riesenprojekt, das von allen Seiten durchdacht werden will. Ein Projektteam mit Mitgliedern aus allen betroffenen Bereichen und Abteilungen muss ran, das sämtliche Aspekte aus der Praxis zusammenträgt, die für eine Entscheidung relevant sind. Da die Poststelle als Eingangs- (und Ausgangs-)Portal eine entscheidende Schnittstelle ist, entsendet auch sie ein Projektmitglied. Der Poststellenleiter wird vortragen, welche Anforderungen an den Scanner zu stellen sind (und dazu das blaue Postmaschinen-Buch mitbringen) und was er von der Capture-Software erwartet. Dafür stellen wir eine Anzahl von Produkten unterschiedlicher Hersteller vor.

Die Vorteile von digitalisierter Post sind in jedem Fall bestechend. Zwar wird das papierlose Büro eine Illusion bleiben – im Gegenteil besteht sogar die Gefahr, dass ein Dokument, das auf den Bildschirm kommt, zu allererst ausgedruckt wird, weil es sich so besser bearbeiten lässt. Aber die gescannte Post lässt sich sehr schnell auch mehrfach und an verschiedene Standorte des Unternehmens selbst weltweit verteilen. Sie archiviert sich automatisch und enorm Platz sparend, und sie kann ohne großen Suchaufwand wiedergefunden werden. Dadurch amortisiert sich die kostspielige Anschaffung schnell. Nicht das papierlose Büro, sondern Unternehmenskommunikation in Echtzeit ist das Ideal.

Capture-Software – das ist die Software, die beim bzw. nach dem Scannen des Posteingangs benötigt wird. Capture bedeutet allgemein in etwa gefangen nehmen, bei der Software trifft „erfassen“ den Vorgang am besten. Sie erfasst das gescannte Image und stellt es entweder in dieser Bildform plus Informationsformular für die weitere Bearbeitung zur Verfügung oder – und das ist die komplexere, höher automatisierte Form – liest die Informationen aus, so dass das Dokument als Volltext-Datei zur Verfügung steht. Diese beiden Arten der Datenextraktion nennt man manuelle und maschinelle Datenerfassung. In diesem Rahmen werden die Daten klassifiziert, interpretiert, mit Datenbanken abgegli-

chen, korrigiert und schließlich in Folgesysteme exportiert, damit die inhaltliche Sachbearbeitung ansetzen kann und die Dokumente archiviert sind.

Hand oder Maschine – auf die Menge kommt es an

Ob es sinnvoller ist, bei der Datenerfassung Hand anzulegen oder sie automatisiert durchzuführen, ist vor allem eine Frage des täglich zu bewältigenden Volumens. Denn je stärker die Prozesse automatisiert werden, desto stärker wächst auch die Komplexität – und damit auch der Preis der Software. Bei der manuellen Datenerfassung wird das Dokument als Bild archiviert. Verknüpft wird es mit einem digitalen Formular, in das einige Indexdaten eingegeben werden müssen. Hierbei liegt der Beleg entweder in Papierform oder auf dem Bildschirm vor. Bei der zweiten Variante ist es möglich, dass nur bestimmte Teile des Belegs angezeigt werden und so eine höhere Sicherheit hinsichtlich Datenschutz gegeben ist. Dabei trägt der Bediener am Bildschirm in das Formular die wichtigen Informationen über das vorliegende Dokument ein: Art des Schreibens (Bestellung, Reklamation, Lieferschein, Bewerbung ...), Kundennummer, Sachbearbeiter und anderes. Dieser Bearbeitungsvorgang wird von der Software unterstützt – das Datum ist schon eingetragen, für die Art des Schreibens gibt es eine Liste zur Auswahl und zur Eindeutigkeit des Eintrags,

über die Kundennummer wird eine Adresse zur leichteren Bearbeitung aus der Datenbank herangeholt, wodurch auch direkt festgestellt werden kann, dass kein Tippfehler vorliegt, sofern diese Adresse mit dem Absender übereinstimmt. Diese Datenerfassung per Hand ist sehr anschaulich, und die Fehlerquote bei der Eingabe ist minimal. Der Capture-Prozess ist jetzt schon abgeschlossen, das digitalisierte Dokument steht der Nachwelt zur Verfügung.

Allerdings ist das nichts für Massenpost. Um für die große Eingangspost nicht ein Heer von zusätzlichen Poststellenmitarbeitern einzusetzen, muss hier eine stärker automatisierte Lösung her. Bei der maschinellen Datenerfassung arbeitet im Hintergrund weitere Software, die die gescannten Bilddaten in Textdaten umwandelt. Die wichtigsten sind:

- OCR – für gedruckten Text
- ICR – für Handschriften
- OMR – für Bildelemente wie Ankreuzkästchen, Feststellen, ob eine Unterschrift vorhanden ist, und Ähnliches.

Solche Software muss nicht bedient werden. Sie hat aber einiges zu leisten, bevor sie ein akzeptables Ergebnis vorlegt, das nur noch wenig Nachbearbeitung erfordert. Für einen Menschen ist es recht trivial, eine Rechnung von einer Bewerbung zu unterscheiden und entsprechend zuzuordnen. An eine Maschine stellt dieser Prozess enorme Anforderungen. Die OCR muss für eine Volltext-Erkennung zum Beispiel grafische Bereiche erkennen und überspringen, statt einen Hieroglyphensalat von unfassbarer Länge zu liefern. Intelligente OCR kann eine Rechnung über Schlüsselwörter wie „Rechnungsnummer“ oder Rech.nr.“ erkennen und sucht in der unmittelbaren Umgebung nach einer Ziffernfolge. Bei Formularen liest sie die Informationen aus festgelegten Feldern aus. Ganz besonders kompliziert sind unstrukturierte Dokumente. Hier ist eine besonders gute Bildqualität

nötig, damit aus dem gescannten Bild ein lesbarer Text wird, den die Software analysieren kann.

Struktur im Kraut – Dokumententypen

Grundsätzlich werden drei Dokumententypen voneinander unterschieden, je nach Grad der Strukturierung – strukturiert, teilstrukturiert, unstrukturiert.



Die Softwares sind bereits so gut in der Klassifizierung, dass die Fehlerquote bei der Dokumentenklassifizierung nicht höher ist als beim Menschen. Der Mensch hat hier also nur wenig zu tun. Nur nach digitalem Achselzucken bei Ausnahmefällen schaut er nach und legt die Dokumentenklasse manuell fest.“

Besonders strukturierte Dokumente sind Formulare. Hier stehen die Informationen an festgelegten Stellen, wodurch sie relativ einfach zu erfassen sind. Ein großer Anteil des Texts wird richtig erkannt, und es ist wenig Nachbearbeitung erforderlich. Allerdings werden viele Formulare mit der Hand ausgefüllt, und vor allem wenn es keine Kästchen für Blockbuchstaben gibt, wachsen Buchstaben oder sogar Zeilen ineinander. Im Extremfall geht die OCR in die Knie.

Teilstrukturierte Dokumente sind in der Regel getippt. Das sind vor allem Bestellungen oder Rechnungen, die immer einem bestimmten Muster folgen und mit regelbasiertem Wissen ausgewertet werden können. Ein Wort wie „Rechnung“ lässt zudem vermuten, dass dem Eingangsbrief eine Bestellung vorausgegangen ist, die im System gefunden und zugeordnet werden kann. Das Erkennungsergebnis bei solchen Dokumenten ist lückenhaft, aber die manuelle Bearbeitung kann deutlich reduziert werden, zumal jede moderne Software lernfähig ist und die Trefferquoten immer weiter ansteigen.

Unstrukturierte Dokumente folgen kaum einem Muster. Hierbei handelt es sich um freie Texte wie Bescheinigungen, Anfragen und Unfallberichte. Kaum maschinell zu lesen sind Handschriften, zumal dort Ober- und Unterlängen der einzelnen Zeilen ineinander greifen und keine klare Zeilenbildung zulassen. Da aber auch Privatpersonen bei ihren Briefen immer mehr dazu übergehen, den

Füller in der Schublade zu lassen und den Rechner hochzufahren, gelangen entsprechend weniger handschriftliche Briefe in den Posteingang. Die meisten unstrukturierten Dokumente können im Volltext erkannt werden – nur bei der Klassifizierung und Indizierung brauchen sie mehr menschliche Unterstützung. In allen Dokumenten können Elemente vorkommen, die bei der Verarbeitung helfen oder stören. Jeweils ein Beispiel: Es gibt Interpretationsmodule, die grafische Elemente nutzen – etwa ein Logo erkennen und so eine Rechnung von einem bestimmten Rechnungsleger identifizieren. Problematisch sind Tabellen. Während sie für das menschliche Gehirn eine hübsche Strukturierung bilden, muss eine Software, die eigentlich von Zeilen ausgeht, sie zunächst als solche erkennen, anschließend die Feldinformationen interpretieren – und dann gibt es auch noch leere Felder.

Schubladen und Fächer – Klassifizierung und Indizierung

Um ein Dokument automatisch an die richtige Stelle weiterzuleiten, muss die Software aus dem Inhalt (bzw. bei einem

Formular aus der Gestaltung) erschließen, um welche Art von Schreiben es sich handelt. In der Regel macht sie sich Schlüsselwörter und Wörterzusammenhänge im Freitext zunutze: „hiermit kündige ich“, „fristgerecht“, eine schlüssige Nummer und eine Adresse deuten unmissverständlich auf Kündigung hin, und zwar mitsamt Vertragspartner und Vertrag.

Die Softwares sind bereits so gut in der Klassifizierung, dass die Fehlerquote bei der Dokumentenklassifizierung nicht höher ist als beim Menschen. Der Mensch hat hier also nur wenig zu tun. Nur nach digitalem Achselzucken bei Ausnahmefällen schaut er nach und legt die Dokumentenklasse manuell fest. Klassifizierung und Indizierung (auch: Indexierung) werden häufig gleichbedeutend verwendet. Wenn nicht, ist die Indizierung die Feinstruktur. Ein Beispiel: Die Dokumentenklasse „Rechnung“ hat die Indexstruktur „Rechnungsnummer“, „Rechnungsleger“, „Rechnungsdatum“ usw.

Alles richtig? – Korrektur und Validierung

Der Erkennungs- und Klassifizierungsvorgang ist blitzschnell abgeschlossen. Jetzt ist wieder menschlicher Einsatz gefragt, denn keine Software kann die Dokumente so perfekt erfassen, dass man sich blind auf das Ergebnis verlassen kann. Korrektur braucht den Menschen, sein Auge auf dem Bildschirm und seine Finger auf der Tastatur. Jedoch leistet die Software Vorarbeit, damit nicht alles abgeglichen werden muss. Könnte sie beispielsweise entweder die Straße oder die Postleitzahl einer Anschrift nicht erkennen, nimmt sie eine Plausibilitätsprüfung vor, bei der sie die Ergebnisse mit einer Adressdatenbank abgleicht und dem Bediener einen Vorschlag unterbreitet, den er nur noch verwerfen oder bestätigen muss. Das ist eine Form der Validierung, ein Abgleich, der im Hintergrund läuft. Es gibt eine Reihe von bewährten Korrekturvarianten auf verschiedenen Ebenen, die sich im Einzelfall mehr oder weniger eignen:

- Um eine *Zeichenkorrektur* zu ermöglichen, zeigt der Bildschirm den gesamten Text an und markiert mit dem Cursor nacheinander die fragwürdigen Zeichen, die der Bediener dann korrekt überschreiben kann.
- Die *Zeichenvvalidierung* ist sehr anstrengend, aber sehr sicher. Hier werden zum Beispiel alle O-Buchstaben angezeigt, worunter eine fehlerhafte 0 (Null) sofort auffällt.
- Bei der *Belegkorrektur* werden der Originalbeleg und der in eine Maske extrahierte Inhalt zum Vergleich nebeneinander auf dem Bildschirm angezeigt. Für den Korrekturingriff gibt es wieder Unterstützung von Datenbanken

und Wörterbüchern, damit nicht alles abgeschrieben werden muss. Sinnvoll ist diese Variante vor allem, wenn der Dokumententyp nicht zugeordnet werden konnte.

- Für die *feldbezogene Korrektur* erscheinen auf dem Monitor nur eine extrahierte Textstelle und die entsprechende Passage aus dem Original. Das ist übersichtlicher und in Bezug auf Datenschutz sehr sicher, da der Bediener nicht das gesamte Dokument zu Gesicht bekommt.
- Die *mehrstufige Korrektur* schließlich bezieht mindestens zwei Korrekturkräfte ein, die beide zum selben Ergebnis kommen müssen, damit der

Korrekturvorgang als abgeschlossen gilt, eventuell bedient sie sich dabei mehrerer Korrekturvarianten. Dieser Aufwand ist sinnvoll bei sehr sensiblen Dokumenten wie Überweisungsträgern.

Die Zeichenerkennung selbst funktioniert bei gedruckten Texten sehr gut; problematisch sind Störungen wie ein Knick mitten durch eine Textzeile oder ein Stempelaufdruck darüber, die Textstellen schwer lesbar machen. Der Aufwand der Nachbereitung hängt also von der Beschaffenheit der Belege ab. Bei guter Qualität steht der Automatisierung eigentlich nichts im Weg, und es ist kaum Nachbereitung nötig.

Die Angaben in der Übersicht

Für die Übersicht haben wir die Hersteller mit einem umfangreichen Fragebogen angeschrieben – die Angaben stammen also aus Herstellerhand und sind trotz redaktioneller Bearbeitung entsprechend heterogen. Es ist deswegen auch möglich, dass Eigenschaften aus einzelnen Einträgen auch auf andere zutreffen, obwohl sie dort nicht erwähnt werden. Da die gesamte Übersicht den größten Teil dieses Hefts einnehmen würde, haben wir sie aufgeteilt. Die Einträge sind alphabetisch nach Hersteller sortiert, in diesem Heft finden Sie lediglich die ersten fünf Hersteller, in den nächsten Ausgaben die Fortsetzung. Aufgrund der Komplexität des Themas „Software für Dokumentenmanagement“ haben wir uns auf den – noch immer hochgradig komplexen – Teil beschränkt, mit dem die Poststelle unmittelbar in Berührung kommt. Ergänzend weisen wir darauf hin, dass es Gesamtlösungen gibt, die nicht allein auf Capturing abheben und sich daher nicht in unser Schema einbinden lassen. Erwähnen wollen wir zum Beispiel Documents for Invoice von Readsoft, das die gesamte Rechnungsbearbeitung übernimmt. Die Software ist in der Übersicht für einen Vergleich dargestellt. Dabei geht

es aber nicht um ein Ranking, denn eine Software ist nicht einfach gut oder schlecht. Vielmehr kommt es auf den Einsatzbereich an. In den einen Unternehmen gehen hauptsächlich Überweisungsträger oder andere Formulare ein, in anderen überwiegend Bestellungen, Lieferscheine und Rechnungen, wieder andere haben es viel mit frei formulierten Schadensmeldungen zu tun. Auch die Menge an Eingangsdokumenten spielt bei der Auswahl eine große Rolle. Eine wichtige Entscheidungsgrundlage, die jedes interessierte Unternehmen selbst klären muss, ist, ob sich ein Produkt in bestehende Systeme integrieren lässt.

Die Angaben der Reihe nach:

Überschrieben sind sie mit dem Namen des Herstellers und dem der aktuellen Software-Version. Es folgt eine kurze Beschreibung, was der User bei der Arbeit mit der Software zu tun hat. Im Anschluss gibt es eine Aufzählung von Eigenschaften:

Allgemein

- 1 Capturing ist ein kleiner, aber wesentlicher Bestandteil des Dokumentenmanagements. Bei einigen Herstellern gehört die Software

- zur eigenen Komplettlösung, bei anderen ist sie ein Modul für Fremdprodukte. In der Aufzählung gibt der erste Punkt Auskunft über Schnittstellen, hieraus geht also der Anbieter der *Komplettlösung* hervor.
- 2 Die Softwares bieten Lösungen für *zentrale und/oder dezentrale Erfassung* (am Arbeitsplatz in der Fachabteilung).
- 3 Gefragt wurde nach einem empfohlenen Richtwert, *wie viele Dokumente am Tag* erfasst werden können. Die meisten Produkte werden beim Einrichten auf kleine bis große Mengen skaliert. Damit ist die Tagesmenge vom Scanner abhängig.
- 4 Bei den *Dokumententypen*, die verarbeitet werden können, standen zur Auswahl: 1. Formulare und andere strukturierte Dokumente, 2. Rechnungen, Lieferscheine und andere teilstrukturierte Dokumente, 3. Bewerbungen, Schadensmeldungen und andere unstrukturierte Dokumente, 4. Druckschriften, 5. Handschriften.
- 5 In der Regel ist eine Capture-Software kein Produkt von der Stange, sondern wird *bedarfsgerecht eingerichtet*. Die Details, inwiefern das die Benutzeroberfläche und die Funktions-

Export: raus aus der Poststelle

Hat die Poststelle die Bearbeitung eines Dokuments abgeschlossen, entlässt sie es auf Klick hinaus in die Welt. Oft sind die Dokumente dann schon im Rahmen der Dokumentenanalyse mit Prioritäten für die Weiterleitung oder weitere Bearbeitung versehen worden.

Die Folgesysteme stammen oft von anderen Herstellern. Bei der Anschaffung ist daher unbedingt darauf zu achten, ob es zwischen diesen beiden Schnittstellen gibt.

Jetzt wandert die Datei ins Archiv, wo sie sich automatisch einsortiert. Gleichzeitig landet sie im Posteingang eines jeden Sachbearbeiters, der sich inhalt-

lich damit befassen soll – eine Bestellung etwa in der Versandabteilung, wo die Waren zusammengestellt werden, und in der Buchhaltung, die die Rechnung dazu schreibt und vermerkt. Doch das ist eine andere Geschichte und soll ein anderes Mal erzählt werden.

Weiterführendes über die nächsten Seiten hinaus

Für alle, die sich mit dem Angebot an DMS-Komplettlösungen befassen wollen oder müssen, ein Tipp: die gigantische VOI-Marktübersicht „Dokumenten-Management-Systeme – Hersteller und Produkte“. Sie ist 2006 erschienen, daher nicht mehr auf dem neuesten

Stand der Versionen. Doch jeder einzelnen der zahlreichen Lösungen sind mehrere Seiten gewidmet, und das in einer Übersichtlichkeit, die unmittelbare Vergleiche zulässt. Doch Vorsicht: Das Werk ist keine Bettlektüre, der dicke Schinken wiegt rund zweieinhalb Kilo. Erhältlich ist es im VOI-Shop unter www.voi.de für 160,50 Euro.

Neu vom VOI gibt es das Buch „Digitale Postbearbeitung – Einführung und Voraussetzungen“ – lesen Sie dazu mehr auf Seite 49.

Die Kontaktdaten der Hersteller aus der nachfolgenden Übersicht finden Sie unter „Capturing“ hinten im Branchenteil dieses Hefts. ak

weise betrifft, klären Sie am besten mit dem Hersteller.

Schnittstellen

- 6 Zu *Hardware-Schnittstellen* haben die Hersteller entweder Twain bzw. Isis genannt oder sich auf den Scanner selbst bezogen, manche gaben sogar eine Empfehlung für einen bestimmten erprobten Scanner.
- 7 Bei den *Schnittstellen zu Folgesystemen* wie Workflow oder Archivierung haben manche Hersteller allgemein geantwortet, andere haben konkrete Angaben gemacht, die sich unter Umständen mit dem ersten Punkt überschneiden.
- 8-10 Das *Einbinden von Dokumenten, die bereits in elektronischer Form vorliegen* (eMail, eMail-Anhänge, Fax-Dateien) ist wichtig, damit bereits beim Eingang, also im frühen Stadium des Dokumentenmanagements, alle Informationen und Eingänge verarbeitet werden können.

Datei und Format

- 11 Der *Dateiname* wird in der Regel automatisch vergeben. Falls dieser Name in der Ablage eine Rolle spielen soll, ist es wichtig zu erfra-

gen, nach welchen Kriterien dies geschieht.

- 12 Die *Art der Erfassung* ist ganz wesentlich für den Grad der automatisierten Arbeit und der anschließenden Ablage. Entweder wird das Image erfasst und zusammen mit einem Formular abgelegt, in dem abrufbare Informationen zum Dokument eingetragen sind, oder aber das Dokument wird im Volltext erfasst und ist komplett elektronisch durchsuchbar, teilweise können auch grafische Elemente mit erfasst werden.
- 13 Bei der *Ausgabe* bzw. dem *Export* von Daten in Folgesysteme geht es entsprechend darum, ob Bild- und Indexdatei bzw. Volltextdaten übergeben werden.
- 14 Auch die möglichen *Dateiformate* korrespondieren mit der Art der Erfassung.
- 15 Die Möglichkeit der *elektronischen Signatur* ist bei vielen Softwares standardmäßig oder optional vorgesehen. Manche Produkte bieten zumindest die einfache digitale Signatur.

Klassifizierung/Indizierung

- 16 Die *Klassifikationsverfahren* sind ein komplexes Thema. Hier unterschei-

den sich die Softwares elementar voneinander. Entsprechend tauchen viele Fachbegriffe auf, die wir hier nicht erschöpfend erläutern können. Vielmehr betrachten wir sie als Grundlage für ein Gespräch zwischen Interessent und Hersteller.

Korrektur/Validierung

- 17 Welche *Korrektur- und Validierungsmöglichkeiten* notwendig sind, ist eine sehr anwenderspezifische Angelegenheit. Gelistet sind hier die Angebote des Produkts.
- 18 Die meisten Softwares können bei der *Kontextanalyse* im Rahmen der Volltexterfassung auf externe Daten zugreifen – beispielsweise Datenbanken mit Adressen oder Wörterbücher.

Referenzen

- 19 Viele Hersteller haben für die Übersicht eine *Liste mit Kunden*, bei denen die Software im Einsatz ist, erstellt.

Weitere Fähigkeiten und Eigenschaften

- 20 Den Abschluss eines Eintrags bilden *Ergänzungen*, die nicht durch den Fragebogen abgedeckt waren.

A2iA: DocumentReader 3.0

Die Images der Dokumente werden von der Software automatisch bearbeitet: Texte, Ziffern, Datum etc. werden in strukturierten, halbstrukturierten und unstrukturierten Dokumenten erkannt und direkt an Folgesysteme geliefert.

- 1 wird in Lösungen von mehr als 200 Partnern integriert, dazu gehören BancTec, Dicom, NCR, Readsoft, Top Image Systems, Prolistic, Wincor-Nixdorf, Saperion, SER, Softpro, I.R.I.S., Cardiff, EMC Captiva
- 2 zentrale Erfassung
- 3 für unbegrenzte Dokumentenmengen
- 4 für alle Dokumentenarten
- 5 wird kundenspezifisch angepasst
- 6 empfohlen wird kein bestimmter Scanner, aber mind. 200 dpi und Graustufen-Bilder
- 7 Schnittstelle ist ein SDK (Software Development Kit), kann zur Schnittstellenprogrammierung in das vorhandene DMS integriert werden
- 8 eMail-Verarbeitung nicht unterstützt
- 9 Verarbeitung von eMail-Anhängen nicht unterstützt
- 10 Fax-Dateien werden nicht verarbeitet
- 11 Vergabe von Dateinamen nicht automatisch
- 12 Volltext-Erfassung: OCR, ICR, IDR und IWR (Intelligent Word Recognition); A2iA hat einen Schwerpunkt auf der Erkennung von Handschriften
- 13 k.A.
- 14 Formate: Tiff, Multi-Tiff, BMP, Jpeg
- 15 elektronische Signatur nicht vorgesehen
- 16 Klassifikationsverfahren:
 - auf Inhaltsbasis durch Extrahieren von Schlüsselwörtern
 - auf Grafikbasis nach Logo und grafischen Zeichen im Dokument
- 17 Validierung/Korrektur: Erkennungsergebnisse werden für jedes Feld bzw. jedes Wort geliefert; mitgelieferte Statistik zur Erkennungsquote
- 18 Kontextanalyse: Grundlage der Handschriftenerkennung ist Abgleich mit Wörterbuch
- 19 Referenzkunden: Société Générale, Deutsche Bank, LBBW, SFR/Arvato-Bertelsmann, Ministerium für Justiz, Ministerium für Finanzen, Ministerium für Verteidigung u.a.
- 20 weitere Fähigkeiten und Eigenschaften:
 - verarbeitet heterogene Dokumentenfluten, z.B. Briefe mit Anlagen wie Formulare, Überweisungen, Kopien von Pässen usw.
 - automatische Vollständigkeitsprüfung eines jeden Vorgangs

Aabby: FlexiCapture 8.0 Professional

Nach dem Scannen oder dem Import aus einem überwachten Verzeichnis werden die Bilder automatisch verarbeitet (Seitentrennung, Template-Matching, OCR/ICR) und ein Wörterbuch- und Datenbankabgleich durchgeführt. Dokumentenbasierte Script-Regeln prüfen die logische Integrität der

erkannten Daten. Die Software visualisiert unsicher erkannte Datenfelder und Zeichen zur manuellen Korrektur. Hierbei sind Kontext- und Gruppenverifizierung möglich. Der Export der geprüften Daten und Dokumente wird mit einem Klick ausgelöst.

- 1 wird über zertifizierte Partner und VARs vertrieben und implementiert; Partner sind in der Regel auf bestimmte DMS spezialisiert
- 2 zentrale und dezentrale Erfassung
- 3 für ca. 100 bis sehr große Dokumentenmengen
- 4 für alle Dokumentenarten in Druckschrift
- 5 wird kundenspezifisch angepasst
- 6 unterstützt Twain und Isis; Dokumente können auch automatisch aus überwachten Verzeichnissen ausgelesen werden (Hot-Folder-Support)
- 7 unterstützt Datei- und ODBC-Export. Durch VB- und Jscript-Unterstützung können eigene Exportroutinen erstellt werden. Microsoft-Sharepoint-Anbindung geplant
- 8 eMail-Verarbeitung indirekt unterstützt
- 9 Verarbeitung von eMail-Anhängen indirekt unterstützt
- 10 Fax-Dateien werden verarbeitet
- 11 automatische Vergabe von Dateinamen; über Regeln steuerbar, auch mit Zugriff auf extrahierte Felder im Dokument
- 12 Erfassung:
 - OCR in 47 Hauptsprachen mit Wörterbuchunterstützung (verschiedene Texttypen – Matrixdrucker, Schreibmaschine, typografisch, OCR-A, OCR-B, MICR (E13B)) und 143 zusätzliche Sprachen mit lateinischen, kyrillischen und griechischen Zeichen
 - ICR in 23 Hauptsprachen mit Morphologie- und Wörterbuchunterstützung, 69 zusätzliche Sprachen mit lateinischen, kyrillischen und griechischen Zeichen
 - OMR: eingekreiste Felder und Zahlen, Auswahlfelder, ungerahmte Auswahlfelder, Kreise
 - Barcode – 1D: Code 39, Check-Code 39, Interleaved 25, EAN 13, EAN 8, Code 128, Codabar, Code 93, IATA 25, UCC-128, UPC-A, UPC-E, Matrix 2 of 5, Industrial 2 of 5, PostNet
 - Barcode – 2D: PDF 417
- 13 Ausgabe: Daten- und Dokumentexport sind möglich über Datei-, ODBC- und scriptbasierten Export
- 14 Formate:
 - Image-Eingabeformate: PDF, PDF/A, BMP, PCX, PNG, Jpeg, Jpeg 2000, Tiff
 - Datenexport: TXT, XLS, DBF, XML, CSV
 - Dokumentenexport: PDF, PDF/A, Tiff, Jpeg, PCX, BMP, PNG, DCX
- 15 elektronische Signatur nicht vorgesehen
- 16 Klassifikationsverfahren: Dokumentenklassifizierung, -sortierung, Lokalisierung von Feldern basierend auf fixen Templates und Flexi-Layouts; dokumentenbasierter Processing-Ansatz, Unterstützung mehrseitiger Dokumente mit sich wiederholenden Elementen, Erkennung und Verarbeitung von Anhängen (Annex Pages)

- 17 Validierung/Korrektur: Gruppenverifizierung, Feld- und Kontextverifizierung, Script-Regeln, „On the fly“ OCR-Korrektur, manuelle Datenerfassung und Korrektur
 - 18 Kontextanalyse mit integrierten Validierungsregeln, Abgleich mit Datenbanken und Prüfsummen, Ersetzen von Werten aus einer Liste, Normalisierung von Daten, verknüpften Wörterbüchern, regulären Ausdrücken, benutzerdefinierten Scripten
 - 19 Referenzkunden auf Anfrage
 - 20 weitere Fähigkeiten und Eigenschaften: Dokumententrennung nach Dateien, Bildern, Leerseiten zwischen verschiedenen Dokumenten, Patchcodes, Barcodes
- 7 Schnittstellen zu DocuWare 5.x, Anpassungen an Import-schnittstellen von gängigen DMS sind möglich
 - 8 eMail-Verarbeitung ist in Planung
 - 9 Verarbeitung von eMail-Anhängen ist in Planung
 - 10 Fax-Dateien werden verarbeitet
 - 11 Dateiname über Export beliebig wählbar
 - 12 manuelle oder vollautomatische Erfassung
 - 13 Ausgabe: Indexdatei und Bilddatei
 - 14 Formate: PCX, Tiff, Multi-Tiff, MO:DCA, PDF, Jpeg, Jpeg2000, Gif, PNG, Plexus-Tiff, BMP, Jbig, CALS u.a.
 - 15 elektronische Signatur ist in Arbeit
 - 16 Klassifikationsverfahren: nach Dokumententypen z.B. durch Barcodes, Formularerkennung, Indexwerte (regelbasiert/grafisch)
 - 17 Validierung/Korrektur: Überprüfung der Felder durch beliebige reguläre Ausdrücke, nachgelagerte Exporte
 - 18 Einbeziehen von Datenbanken, Wörterbüchern etc. in die Kontextanalyse in Vorbereitung
 - 19 Referenzkunden: Fiducia IT, Telekom Vivento, mehrere Dienstleister, SSC
 - 20 weitere Fähigkeiten und Eigenschaften: Erweiterung des Systems auch auf weitere Arbeitsplätze problemlos möglich; Einbinden eigener Applikationen möglich; mehrere Stapel können gleichzeitig bearbeitet werden

Alos: Alos-Scan 5.0

Die Benutzer müssen nur aus der Stapelverarbeitung Arbeitsschritte auswählen, die manuelle Tätigkeiten erfordern. Der Rest wird automatisch durchgeführt.

- 1 DocuWare 5.x
- 2 zentrale und dezentrale Erfassung
- 3 bis zu 200.000 Belege pro Arbeitsplatz und Tag
- 4 für alle Dokumentenarten in Druckschrift
- 5 wird kundenspezifisch angepasst
- 6 unterstützt Isis-Scanner und Kofax-zertifizierte Scanner

Unser Mandant ist ein privates Briefdienstleistungsunternehmen und gehört zu einer erfolgreichen Mediengruppe in Süddeutschland. In dem sich stark verändernden Briefmarkt ist das Wachstum enorm.

Deshalb suchen wir für den weiteren Ausbau der Marktanteile eine(n)

Geschäftsführer/in Briefzustellung

In dieser Position verantworten und koordinieren Sie sämtliche Aktivitäten im Bereich der privaten Briefzustellung. Dabei bilden folgende Aufgaben den Schwerpunkt Ihrer Tätigkeit:

- **Gesamtverantwortung für alle Verkaufsaktivitäten inkl. Akquise und Betreuung der Key-Accounts**
- **Organisatorische und inhaltliche Weiterentwicklung des Unternehmens**
- **Gesamtverantwortung für die Dienstleistungserbringung und das Qualitätsmanagement**
- **Dauerhafte Optimierung von Kosten-, Prozess- und Arbeitsorganisationsstrukturen**
- **Budget- und Investitionsplanung und das zugehörige Controlling**

Sie haben nach Ihrem wirtschaftswissenschaftlichen Studium bereits einschlägige Berufserfahrung in der Logistik / Zustellbranche, idealerweise in der privaten Briefzustellung, z.B. als Stellvertretende/r Geschäftsführer/in oder als Assistent/in der Geschäftsführung gesammelt und möchten nun den nächsten Schritt auf der Karriereleiter in einer spannenden, zukunftsorientierten Branche gehen.

Ihre vollständigen Bewerbungsunterlagen senden Sie bitte mit Angabe Ihres frühestmöglichen Eintrittstermins und Ihrer Gehaltsvorstellungen z. Hd. Herrn Bernd Kröger an u.a. Anschrift oder per E-Mail an kroeger@tgmc.de. Ihre Sperrvermerke werden strikt beachtet.



AnyDoc Software: OCR für AnyDoc 5.0

Die Software arbeitet vollkommen automatisch. Auf Wunsch können auch Interventionen eingebaut werden:

- automatischer Import ins System
 - Qualitätssicherung der gescannten Belege
 - Verifikation, dokument- oder stapelweise
 - automatische Übergabe im definierten Format
- 1 Schnittstellen zu DM-Systemen von Open Text, Hummingbird, EMC, DocuWare, IBM, Hyland, FileNet, Microsoft, Scan View, Windream, Easy
 - 2 zentrale und dezentrale Erfassung
 - 3 für ca. 100 bis sehr große Dokumentenmengen
 - 4 für alle Dokumentenarten
 - 5 wird kundenspezifisch angepasst
 - 6 für alle Twain-, Isis- und Opex-kompatiblen Scanner
 - 7 Schnittstellen zu ODBC-Datenbanken (Oracle, Sequel-Server, SQL-Server), EDI, XML, SAP, BAAN, Oracle Financial, Navision sowie anderen gängigen Folgesystemen
 - 8 eMails werden verarbeitet
 - 9 eMail-Anhänge werden verarbeitet
 - 10 Fax-Dateien werden verarbeitet
 - 11 automatische Vergabe von Dateinamen beliebig definierbar, z.B. ID, Datum, Zeit, Speicherort, Verantwortliche, können auch bestehende Namen von externen Applikationen sein
 - 12 Erfassung: Volltext OCR, Zonal OCR, ICR, OMR und manuelle Indexierung
 - 13 Ausgabe: Volltext-, Bild- (Schwarzweiß/Farbe), Indexdatei möglich
 - 14 Formate: Tiff, Multi-Tiff, Jpeg, PDF, PDF/A, Gif
 - 15 elektronische Signatur optional
 - 16 Klassifikationsverfahren: regelbasiert, statistisch, geometrisch, grafisch, Volltext und Data-Matching mit externen DB-Werten
 - 17 Validierung/Korrektur: Belegkorrektur, feldbezogene Korrektur, Zeichenkorrektur, Batch-Balancing, Doppelverifikation (4-Augen-Prinzip), Supervisor-Verifikation
 - 18 Datenbanken, Wörterbücher etc. werden bei Kontextanalyse eingebunden
 - 19 Referenzkunden: Air Liquide, Arvato Media, Bühler, Capita, Coop, Schweizerische Post, Stadt Luzern, Deutsche Post In Haus Service, div. Schweizer Banken, Holcim, Hotelplan, Leaseplan, Lufthansa A.S., Novartis Pharma, Phonak, PrimaCon Management, SAS Regional Spital, SR Technics, Volkswagen O.T.L., W.L. Gore
 - 20 weitere Fähigkeiten und Eigenschaften:
 - echte Freiform-Erfassung
 - alle Dokumentenverarbeitungsphasen in einem System integriert
 - hohe Dokumenten- und Datensicherheit während der Verarbeitung
 - modulare Architektur, beliebig skalierbar
 - alle Erfassungs-Technologien in einem System integriert
 - optional: Qualitätssicherung der gescannten Belege

Captaris: DOKuStar Professional 3.0

Nach dem Scannen können optional kein, ein oder zwei Validierungsschritte folgen. Beim Zwei-Schritt-Verfahren werden im ersten Schritt die Dokumentenstruktur und die Klassifikation validiert, also ggf. die Dokumentenart geändert oder ein Dokument in mehrere einzelne aufgeteilt, und im zweiten Schritt werden die Indexdaten validiert. Beim Ein-Schritt-Verfahren werden die beiden Arbeiten in einem Schritt erledigt. Welches Verfahren angewandt wird, kann je nach Dokumentenart eingestellt werden.

- 1 ist ein reines Erfassungsprodukt
- 2 zentrale und dezentrale Erfassung: zentral scannen, klassifizieren und indizieren. Nutzdaten, die in den Fachabteilungen benötigt werden, können entweder zentral und vollautomatisch oder in den Fachabteilungen mit Single Click Entry halbautomatisch erfasst werden
- 3 Tagesmengen nach oben hin unbegrenzt, nach unten technisch ebenfalls, wirtschaftlich ab 200 bis 500 Dokumenten pro Tag
- 4 für alle Dokumentenarten
- 5 wird kundenspezifisch angepasst
- 6 für alle gängigen Scanner
- 7 Schnittstellen zu den meisten größeren ECM-Systemen sowie zu Fachanwendungen wie SAP
- 8 eMails werden verarbeitet
- 9 eMail-Anhänge werden über Customizing verarbeitet
- 10 Fax-Dateien werden verarbeitet
- 11 Dateiname wird automatisch generiert, das Schema kann vorgegeben werden
- 12 manuelle, halbautomatische (Single Click Entry) und vollautomatische Erfassung
- 13 Ausgabe: Indexdatei plus Bilddatei sowie Volltextdatei möglich
- 14 alle gängigen Formate möglich, empfohlen wird Jpeg
- 15 elektronische Signatur ist vorgesehen
- 16 Klassifikationsverfahren: unterstützt werden regelbasierte, statistische, geometrische und grafische Verfahren, am häufigsten genutzt werden regelbasierte und statistische. Bei der regelbasierten Klassifikation können grafische Merkmale genutzt werden
- 17 Validierung/Korrektur: Belegkorrektur ergibt sich zwangsweise, möglich auch feldbezogene und Zeichenkorrektur
- 18 Datenbanken, Wörterbücher etc. werden in die Kontextanalyse eingebunden
- 19 Referenzkunden: T-Com, diverse AOK, Landesamt für Besoldung und Versorgung, Swisscom, Interserv u.a.
- 20 Weitere Fähigkeiten und Eigenschaften: Die automatische Erkennung lässt sich über Programmierschnittstellen an unterschiedliche Aufgaben anpassen. Für Anwendungen wie Rechnungsbearbeitung werden vorgefertigte Module angeboten. Mit der semi-automatischen Erfassung Single Click Entry lassen sich auf einfachste Weise unterschiedliche Zielsysteme einbinden.