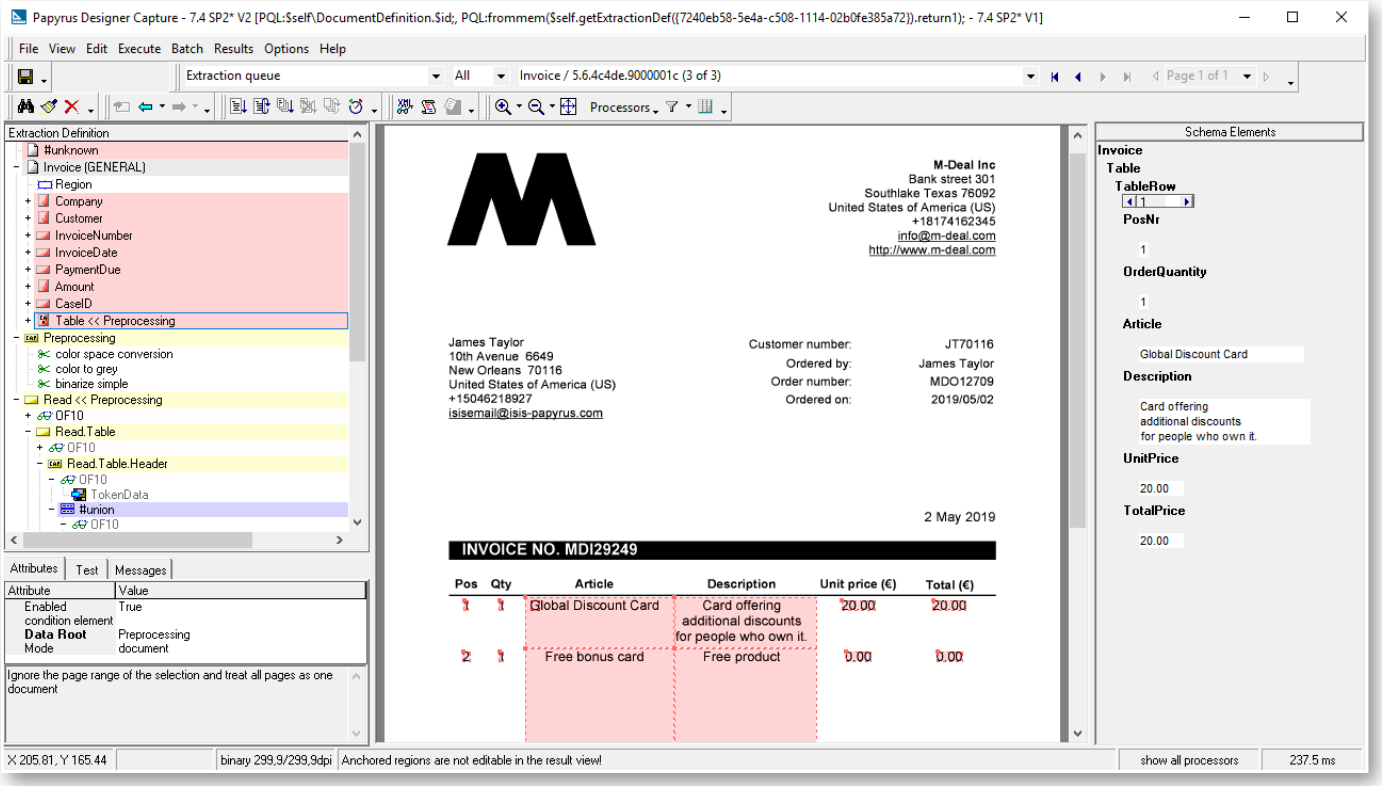
Papyrus Designer/Capture



Produktbeschreibung

Der Papyrus Designer/Capture ist ein sehr leistungsfähiges Werkzeug zur Definition von Klassifizierungs- und Datenextraktionsanwendungen für die automatisierte Verarbeitung aller Arten von Dokumenten.

Papyrus Designer/Extract ermöglicht die einfachen Erstellung von Extraktionsdefinitionen für gescannten oder elektronische unsortierte Geschäftsdokumente mit unbekannter Struktur und Layout wie Rechnungen, Listen und komplexe Tabellen. Es bietet auch alle notwendigen Funktionen für eine rasche Datenextraktionsdefinition zur Verarbeitung aller Arten von strukturierten Formularen.

Papyrus Designer/Classify umfasst ein selbstlernendes Modul zur Klassifizierung von Dokumenten für ein breites Anwendungsspektrum im Bereich der automatisierten Erkennung, Sortierung und Verteilung von elektronischen Dokumenten, Fax als auch Papierpost.

Papyrus Designer/Capture ist somit mehr als ein reines Designwerkzeug zur Texterkennung und erfüllt die Herausforderungen der modernen Geschäftswelt, nur die geschäftsrelevanten Inhalte aus hochvolumiger Papier- und digitaler Korrespondenz zu extrahieren. Das besondere ist die exzellente Fähigkeit zur Identifizierung der relevanten Geschäftsinhalte durch eine Kombination von Mustererkennung, Schlüsselwörter, relativen Positionen, Regeln etc.

Unübertroffene Leistung zur Klassifikation und Datenextraktion aus allen Dokumententypen

Kontakt

Europa
P: +43-2236-27551
F: +43-2236-21081

Amerika
P: 817-416-2345
F: 817-416-1223

Asien Pazifik
P: +65-6339-8719
F: +65-6336-6933

E-Mail & Web
info@isis-papyrus.com
www.isis-papyrus.com

© 2020, ISIS Papyrus, ersetzt bisherige Dokumentation.

Hauptvorteile

- Digitalisierung von Inhalten: Extraktionsdefinitionen von geschäftsrelevanten Inhalten aus hochvolumigen Posteingangsdaten
- Klassifikations- und Extraktionsdefinitionen sowohl für FixForm als auch für FreeForm®-Dokumente
- Unterstützung von Dokumenten üblicher Eingangskanäle in allen Dateiformaten
- Leistungsstarke Merkmale von Klassifikations- und Extraktionsdefinitionen
- Unterstützung verschiedenster Erkennungsalgorithmen für alle Typen von Inhalten
- In Kombination mit Papyrus Recognition Server wird ein hohes Maß an Automatisation zur Digitalisierung aller Inhalte erreicht
- Schnelle Definition und Prüfung von neuen Dokumenttypen

Features

Komponenten

- Papyrus Designer/Extract
- Papyrus Designer/Classify

Imagevorverarbeitung

Papyrus Designer Capture ermöglicht automatische oder angepasste Vorbereitung der Bilddaten für eine optimale Image Verarbeitung. Dies ist besonders wichtig bei gescannten Bildern und Fotos.

- Entfernung von Schmutz und grauen Schatten
- Ringmappenloch-Entfernung
- Automatische Drehung
- Binarisierung
- Farbe in Grau umwandeln
- Kästchen-, Linien-, Gitter- und Kamm-Entfernung
- Blindfarbe
- Verstärkung schwarzer Farbe
- Reduktion von schwarzer Farbe
- Entzerrung der Ränder
- Entzerrung des Inhalts

Erkennungsalgorithmen

Der Inhalt vorverarbeiteter Bilder wird durch einen leistungsfähigen Satz von Algorithmen mit folgenden Eigenschaften bearbeitet:

- Optische Zeichenerkennung (OCR) - maschinell geschriebener Text
- Intelligente Zeichenerkennung (ICR) - handgeschriebener Text
- Optische Markierungserkennung (OMR) - Kontrollkästchen
- Strichcode, QR-Code, DataMatrix
- Logo

Papyrus-Designer/Classify

Diese Komponente wird zur Definition von Kategorien, der Extraktion von Merkmalssätzen und der Klassifizierungsstrategie verwendet.

Mehrerer Definitionsschritte können automatisiert werden zusammen mit einer einfachen Handhabung von Beispieldaten. Papyrus-Designer/Classify ist vollständig in Papyrus Objects integriert.

Klassifizierungsschritte basieren auf:

- Schlüsselwörter und Kombinationen
- Layouts
- Fortgeschrittene Textanalyse
- Logos

Zuerst definieren Sie die erforderlichen Dokumenttypen. Mit einer Anzahl von repräsentativen Stichproben (einige Dutzend) für jede Kategorie wird das System auf die spezifischen Eigenschaften der einzelnen Dokumentenklassen trainiert. Neue Kategorien können einfach und bequem hinzugefügt werden.

Papyrus-Designer/Classify ermöglicht Klassifikation über Regeln, die auch erst während der Extraktion definiert werden können.

Papyrus-Designer/Classify

Um relevante Inhalte zu erkennen, stehen folgende Möglichkeiten und Kombinationen für Extraktionsdefinitionen zur Verfügung

- Muster
- Schlüsselwörter
- Regionen
- Positionen
- Ankerpunkte
- Regeln
- Layout
- Logo

Grundlegender Definitionsablauf

- Definieren Sie alle vorkommenden Dokumentarten
- Definieren Sie die interessanten Elemente für jeden Dokumententyp
- Definieren Sie für jedes Datenelement:
- Allgemeine Attribute
- Verwandte Anker
- Muster, die auf das Element zutreffen
- Bedingungen (Regeln), die erfüllt sein müssen
- Testen Sie die Definitionen mit verschiedenen Dokumenten

Vergleichen Sie die Ergebnisse mit früher generierten "goldenen Daten" Verschiedene Anzeigemöglichkeiten um rasch neue Dokumente zu definieren stehen zur Verfügung:

- Mögliche Extraktionsfelder anzeigen
- Regionen und Zusammenhänge anzeigen
- Bequeme Tag-Werkzeuge

Papyrus Capture/Extract bietet die Dokumentenstruktur tiefgreifend zu analysieren um eine einfache Neuklassifikation zu erreichen. Dies bietet eine große Flexibilität für die Dokumentengestaltung und -verarbeitung von komplexen Dokumenten.

Datenformate

- TIFF 6.0 (auch Mehrfachseiten)
- JPEG, JPEG 2000, BMP, GIF, PNG
- HTML, CSV, TXT (einfacher Text)
- PDF 1.4 oder höher, PDF/A (ISO 19005)
- AFP
- Office Formate:
  - DOC/DOCX/DOCM/ODT
  - PPT/PPTX/PPTM/ODP
  - XLS/XLSX/XLSM/ODS

Voraussetzungen

- Windows Vista/7/8/10/Server 2008/Server 2012/Server 2016
- Standard Workstation
- Min. Bildschirmauflösung 1280x1024
- Papyrus WebRepository

Bestellinformation

Papyrus Designer/Capture
für Windows

Training

ISIS Papyrus bietet seinen Kunden ein spezielles Inhouse Schulungsservice sowie Standardkurse an.

Softwarewartung

Es gelten die AGB für Papyrus Softwareprodukte. Die kostenlose Servicedauer beträgt 6 Monate ab Installation. Danach wird ein Wartungsvertrag optional angeboten.