



Papyrus Designer Package/Capture

The screenshot shows the PapyrusCapture / FreeForm Designer software interface. It features a menu bar (File, Edit, View, Sample, Recognition, Help), a toolbar, and a main workspace. On the left, there is a 'Class' tree showing a hierarchy of document elements like Invoice, Invoicedate, Invoicenummer, etc. Below this is a 'Patterns' table with columns for 'U', 'Express', 'Ignore', 'O', and 'Prior'. The main workspace displays a scanned invoice document. On the right, there is a 'Training' panel with a 'Result' table showing extracted data for the invoice, such as Invoicedate (24.01.2003), Invoicenummer (319019199), and various amounts. Five callout boxes with arrows point to specific features: 'Dokumenten & Feld Struktur' points to the class tree; '5 OCR/ICR engines' points to the Patterns table; 'Definieren der Erkennungsparameter' points to the 'U' column in the Patterns table; 'Anzeige der Testdokumente' points to the main workspace; and 'Direkte Ansicht der Erkennungsergebnisse' points to the Training Result table.

Herausragende Leistungsfähigkeit zur Definition von Klassifikation und Datenextraktion für alle Arten von Dokumenten

Produktbeschreibung

Designer Familie zur Definition von Projekten zur Klassifikation und Datenauslesung für jegliche Art von Dokumenten. Als wesentliche Eigenschaften zur Vereinfachung und Beschleunigung des Customizing-Prozesses sind u.a. auch die beliebige Wiederverwendbarkeit von vordefinierten Parameter-Sets, wo immer möglich „Teach-by-Example“ und selbsterklärende Regeln beinhaltet.

Papyrus Capture Designer/FreeForm® ist ein besonders mächtiges Werkzeug zur einfachen Erstellung von Definitionen für die automatische Datenerkennung aus gescannten, aber unsortierten Geschäftsdokumenten mit unbekannter Struktur und Layout wie z.B. Rechnungen, Listen, formlose Schreiben, etc.

Papyrus Capture Designer/FixForm bietet alle nötigen Werkzeuge für die rasche Definition von Datenextraktionsparametern zur Verarbeitung aller Arten von Formularen mit bekanntem Layout einschließlich Handschrift.

Papyrus Capture Designer/Classify beinhaltet selbstlernende Module zur Klassifikation von beliebigen Dokumenten und ermöglicht dadurch vielfältige Anwendungen im Bereich der automatischen Sortierung und Verteilung von elektronischen Dokumenten, Faxen und Papier-Post.

Kontakt

Europa

T: +43-2236-27551
F: +43-2236-21081

Amerika

T: 817-416-2345
F: 817-416-1223

Asien/Pazifik

T: +65-6339-8719
F: +65-6336-6933

E-Mail

info@isis-papyrus.com

Internet

www.isis-papyrus.com

© 2012, ISIS Papyrus, ersetzt alle bisherige Dokumentation.

1. Jänner 2012

Vorteile

- Schnellere Definition und Testen von neuen Dokumententypen als bei vielen anderen Produkten
- Je ein spezielles Design-Werkzeug für die drei grundsätzlichen Anforderungen des „Dokument-Verstehens“: Klassifikation, Datenauslesen FixForm und FreeForm®
- Ergonomische, selbsterklärende Benutzeroberfläche – keine Programmierkenntnisse erforderlich
- Universelle Repository-Funktionen ermöglichen Wiederverwendung von bereits erstellten Definitionsteilen
- Analyse- und Statistik-Funktionen (Regression Tests) inkludiert

Komponenten des Pakets

Papyrus Designer/FreeForm®

Papyrus Designer/FixForm

Papyrus Designer/Classify

Gemeinsame Eigenschaften

Bildvorverarbeitung

Alle Designer ermöglichen automatische oder einstellbare Aufbereitung der Bilddaten für optimale Erkennungsergebnisse:

- Schmutzentfernung
- Auffetten
- Linienentfernung
- Entfernung von Lochungen
- Automatisches Aufrichten
- Binarisierung
- Umwandlung Farbe in Grau
- Geraderichten mittels
 - Blattrand
 - Textzeilen

Image Erkennung von Ankern, Bitmaps, Linien; Barcode, OMR, Text: 5 OCR/ICR Engines decken jede Erkennungsherausforderung ab, der jeweilige Zeichentyp ist leicht einstellbar.

Flexible Post-OCR-Nachverarbeitung mit Filtermuster, Abstandsunterdrückung u.v.a.m.

Designer/FreeForm®

Der „Star“ unter den Capture Designern ermöglicht sowohl Dokumententyp-basierte generische als auch Template-basierte Extraktion. Die Kombination der beiden Methoden ergibt kürzeste Definitionszeiten und beste Erkennungsergebnisse.

- Generische Definition
 - Regelbasierte Extraktion
 - Wartbares Regelsystem
 - Vorteil: funktioniert auch mit unbekanntem Dokumentlayouts

- Template-basierte Definition
 - Learn-by-example
 - an Dokumentregionen orientierte Datenextraktion
 - Optimierbar auf spezifische Eigenschaften von Dokumententypen.

Struktur des Definitionsprozesses:

- Definition aller auftretenden Dokumententypen
- Definition aller benötigten Datenfelder
- Definitionen für jedes Datenfeld:
 - Allgemeine Eigenschaften
 - Dazugehörige Anker
 - Mögliche Muster und Strukturen des Elements
 - Bedingungen, die erfüllt sein müssen
- Test der Definition mit verschiedenen Dokumenten
- Vergleich der Ergebnisse mit zuvor generierten Soll-Resultaten („Golden Files“)

Tools zur Anzeigeunterstützung ermöglichen die Definition neuer Dokumententypen in Minuten anstatt Stunden:

- Anzeige von Extraktionskandidaten
- Anzeige von Regionen und Abhängigkeiten
- Bequeme Auswahlwerkzeuge.

Papyrus Designer/FixForm

Durchgängig angelegt gemäß XML Struktur-Prinzipien, vereint der Designer/FixForm mehr als ein Jahrzehnt Erfahrung mit Formularauslesung.

Der übersichtliche Arbeitsplatz umfaßt alle nötigen Teilbereiche

- Image-frame zur Anzeige des auszu-lesenden Dokumentes
- Extraktionsparameter wie Lesezone oder Erkennungstyp
- Eigenschaftsfenster mit detaillierten Attributen jedes ausgewählten Elements
- Bereich mit Anzeige der Ergebnisse zur sofortigen Visualisierung von Bildbearbeitung oder erkannten Zeichen.

Papyrus Designer/Classify

Der Designer/Classify dient zur Definition der verschiedenen Dokumentenkategorien, der Extraktion der Feature-Sets und der Klassifikationsstrategie. Er unterstützt die Automatisierung einzelner Definitionsschritte. Der Designer/Classify, weitgehend integriert in die Papyrus Objects Infrastruktur, ermöglicht den einfachen Umgang mit Sample Daten, Testläufen und Ergebnisanalyse.

Unterstützte Klassifikationsarten:

- Image Attribute der ganzen Seite oder einer speziellen Region
- Schlüsselworte auf vorbestimmten Positionen oder an jeder beliebigen Stelle im Text
- Text-Phrasen auf vorbestimmten Positionen oder irgendwo im Text
- Regeln.

Der Designer/Classify visualisiert die automatisch generierten Kategorisierungs-Attribute und ermöglicht

- die Anpassung von Parametern wie die Relevanz von Schlüsselwörtern oder Klassifikations-Schwellwerte
- die Analyse der Erkennungs-, Rückweisungs- und Fehlerraten und auf welchen Dokumenten diese auftreten
- die Adaptierung von Klassifikationsmethoden für spezielle Dokumententypen.

Vor Beginn der Arbeit mit dem Designer/Classify müssen die gewünschten Dokumentenkategorien festgelegt werden. Mit einer ausreichenden Anzahl von Musterdokumenten (einige Dutzend) für jede Kategorie wird das System antrainiert, um die spezifischen Eigenschaften jeder Klasse zu lernen. Neue Kategorien können auch während der Produktionsphase leicht hinzugefügt werden.

Kontinuierliches Zusatztraining als unmittelbares Ergebnis der manuellen Klassifikation ist ebenfalls möglich.

Datenformate

Eine Anzahl von Dateiformaten, Image (TIFF, JPG, PDF) und Text (DOS, ANSI, Unicode), Microsoft Office Files wird unterstützt.

Technische Anforderungen

Hardware: Pentium >1GHz, 1GB RAM, 2GB HD

Bestellinformation

Papyrus Designer Package /Capture 7 für Windows 2K/XP/2K3/VISTA/2K8/7

Schulungen

ISIS bietet seinen Kunden spezielle inhouse trainings services, die *Papyrus Designer Capture Workshops*. Bitte kontaktieren Sie uns bezüglich Einzelheiten und Terminen.

Softwarewartung

Es gelten die Allgemeinen Geschäftsbedingungen für ISIS Software-Produkte. Freie Serviceperiode ist 6 Monate nach Installation. Wartung und Service ist danach optional für Updates und Hotline verfügbar