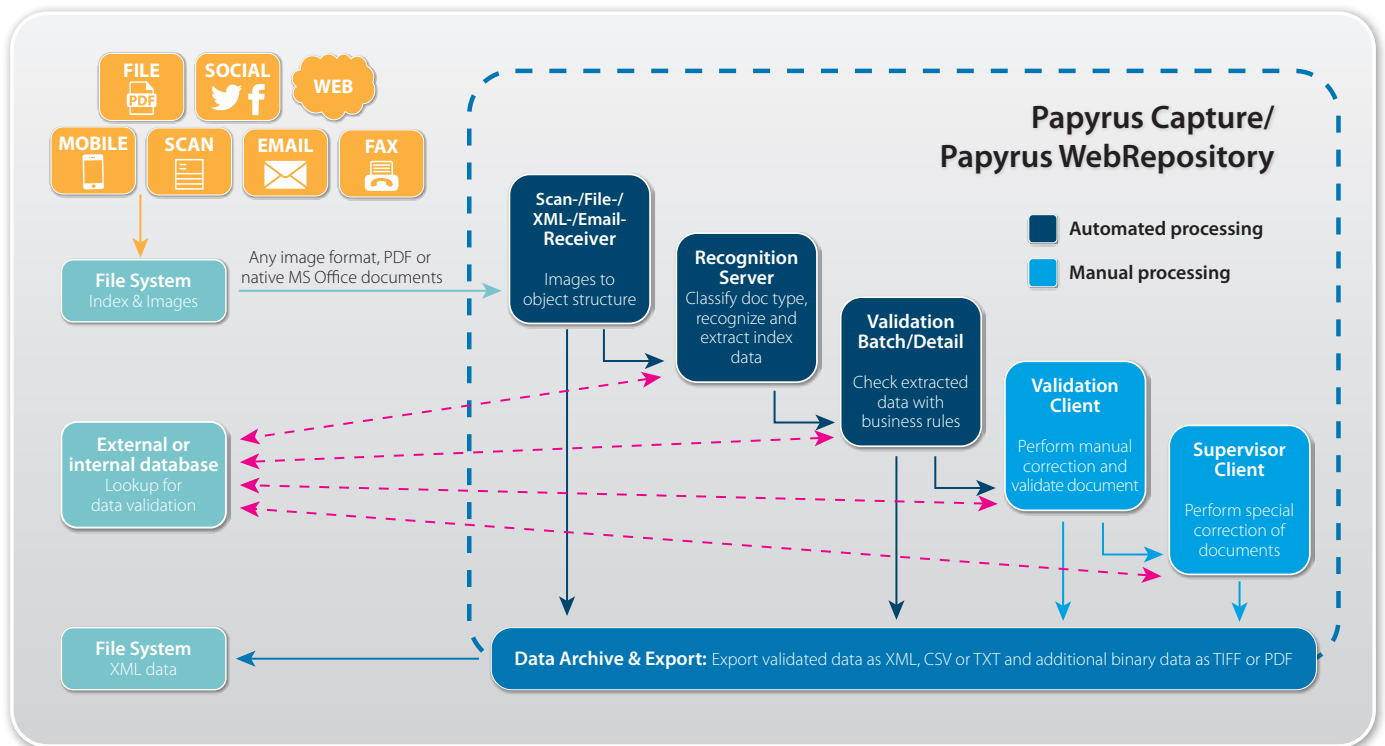


Papyrus Document Capture System 7



Concept Outline

Every day a company receives a multitude of different documents - letters, image data, PDF and Office documents - all of which must be processed quickly and forwarded to the right place. Therefore, the configuration and operation of a capture system must be fast and efficient, regardless of document types, capture hardware and operating systems used.

Based on the Papyrus Business Application Platform, ISIS Papyrus has developed the most powerful tools for building document capture applications. Large banks, insurance companies, government organizations, industry and service companies have one thing in common, they need fast development, elimination of programming effort and absolute independence from operating systems and scanners - all inherent to the **Papyrus Document Capture System**.

The Papyrus Document Capture concept provides an integrated corporate solution for powerful and efficient development, operation and management of document capture systems.

Contact

Europe

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

America

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

Asia Pacific

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

Email & Web

info@isis-papyrus.com
www.isis-papyrus.com

© 2020, ISIS Papyrus,
replaces all previous
documentation.

The Capture Applications

The ISIS Papyrus Document Capture System covers the entire spectrum of possible inbound applications: Scanning or importing documents, classifying and extracting indexes and data, post-processing the results and transferring documents and data to any business application. This includes the recognition of document layouts as well as the extraction of specific contents under consideration of definable document logic, interfaces for data validation and highly functional masks for the completion of the data.

Since all recognition engines are already available during design time, capture quality and results can be verified right from the beginning. Additional tools for the high-volume benchmarking of detection and rejection rates guarantee optimal preparation before going live.

Additional tools for high volume benchmarking of automation and rejection rates guarantee optimal production preparation before going live. These frameworks include all necessary document and data field definitions as well as ready-made views for business administration and workflow design.

Papyrus Capture Products

Papyrus Capture Recognition Server for Windows and Linux

Papyrus Recognition Server is the technological basis for the development of self-learning modules to classify business documents and extract their data. Recognition Server offers a wide range of document capture and extraction methods for structured, semi-structured and completely unstructured documents. It also includes automated sorting and distribution of electronic documents (PDF, Office formats, image formats), fax and paper mail.

Papyrus Capture Designer for Windows

The Papyrus Capture Designer is an application used by the Capture application developer for the design and development of elaborate data extraction projects. It features FixForm and FreeForm® technologies which enable the creation of extraction definitions for data from structured, semi-structured as well as unstructured documents. Its sophisticated, user-friendly interface enables the smooth development of hierarchically organized definitions for the extraction of document fields.

Papyrus Business Designer/Capture for Windows

The Papyrus Business Designer/Capture is the core interface allowing business users to work with the document capture solution. It employs user-trained algorithms to continuously improve the capture results of the powerful OCR engines (for machine writing) and ICR engines (for hand writing) during content and case management operations, such as ad-hoc and on-the-fly extraction/indexing of documents in a business case. This is possible because the Papyrus Business Designer/Capture does not require IT-involvement but gives all the power into the hands of the actual business user.

Papyrus Client/Capture for Windows

The Papyrus Client/Capture is the interface for all kinds of user interactions within a document capture system. It enables, among other things, the completion and correction of classification and extraction results with all necessary display and validation functions.

Papyrus Scan for Windows

Papyrus Scan is the Papyrus WebRepository-based application for scanning documents and automatic import of scanned images into the Depot Nodes of the Papyrus WebRepository. It supports Twain and EMC Captiva ISIS PixTools scanner interfaces and can be used with both document and barcode scanners.

Papyrus WebRepository for z/OS Unix, z/Linux, HP-UX Itanium, AIX, Solaris, Linux and Windows

The Papyrus WebRepository is the core component of the Papyrus Objects Business Application Platform. It enables the control of all applications and document types across all platforms and output channels, for both online and offline users. It provides customizable audit and reporting capabilities and Papyrus Adapters (e.g. Scan, Fax, Email, SOAP, REST) enable seamless integration with other business applications.

Papyrus WebArchive for z/OS native, z/OS Unix, z/Linux, HP-UX Itanium, AIX, Solaris, Linux and Windows

The Papyrus WebArchive offers an integrated archiving solution consisting of a central or distributed short-term and/or long-term archive for all incoming and outgoing document types in any format. Optionally, the storage of associated business cases is also possible. It offers a document query and retrieval user interface that integrates with other Papyrus business applications via the Papyrus EYE Widgets technology, a web browser or the Papyrus Enterprise Mobile App for iOS and Android devices.

Prerequisites

Windows Vista/7/8/10/Server 2008/Server 2012/Server 2016
Papyrus Capture Recognition Server is also available for Linux.

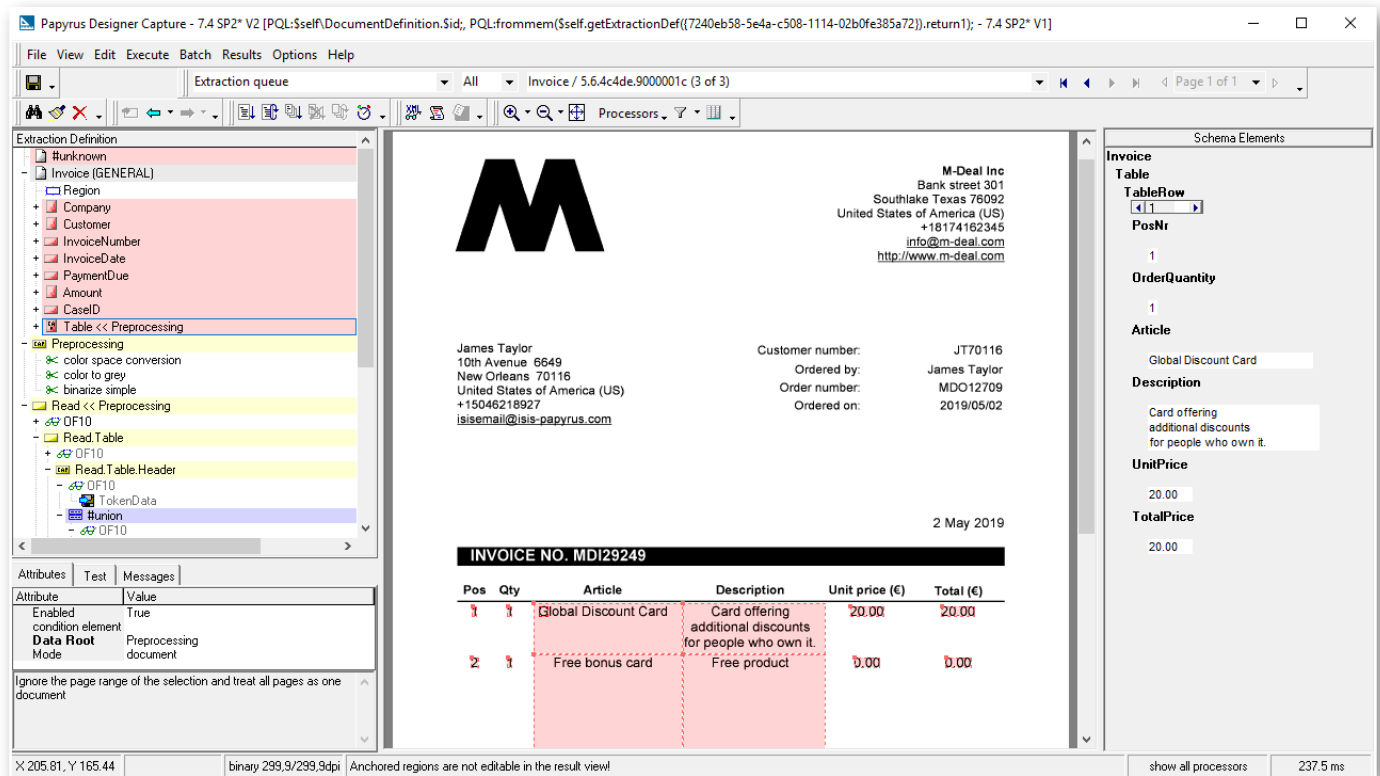
Order Information

The ISIS Papyrus Document Capture System consists of independent components. Even though these are completely integrated, most of them can be used standalone if required.

Maintenance

The Terms and Conditions for ISIS Papyrus Software Products apply. Free service period is 6 months after installation. A maintenance agreement is offered optionally for updates and hotline.

Papyrus Designer/Capture



Product Description

The **Papyrus Designer/Capture** is a very powerful toolkit to define classification and data extraction setups for automated data recognition of different kinds of documents.

Papyrus Designer/Extract component helps to easily create definitions of scanned or electronic unsorted business documents with unknown structure and layouts such as invoices, lists and complex tables. It also provides all necessary tools for a prompt definition of data extraction parameters to process all kinds of structured form layouts.

Papyrus Designer/Classify component supports a self-learning module for classification of documents that offers a broad range of possibilities for application in the field of automated sorting and distribution of electronic documents, fax and paper mail.

Accepting a challenge of the modern world to extract only business relevant content out of high volume inbound paper and digital correspondence, the **Papyrus Designer/Capture** is much more than just a text recognition design tool. It has an excellent capability to identify required business content by means of pattern recognition, keywords, relative positions, rules etc. and a powerful mixture of these means.

Unmatched power to define classification and data extraction on all types of documents.

Contact

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

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

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

Email & Web
info@isis-papyrus.com
www.isis-papyrus.com

© 2020, ISIS Papyrus,
replaces all previous
documentation.

Key Benefits

- Content digitizing: extraction definition of business relevant content out of high volumes of inbound data
- Classification and extraction definition for both FixForm and FreeForm® documents
- Support of documents from all common inbound channels and file formats
- Powerful set of classification and extraction definition features
- Support of diverse recognition engines for different kinds of content
- In collaboration with Papyrus Recognition Server, provides a high-level automated setup for content digitizing
- Fast definition and testing of new document types

Features

Package Components

- Papyrus Designer/Extract
- Papyrus Designer/Classify

Image Preprocessing

Papyrus Designer Capture allows automatic or adjusted preparation of the image data for optimal recognition results. This is especially important for scanned and photo images.

- Dirt removal, Despeckle
- Punch hole removal
- Automatic rotation
- Binarize
- Convert color to gray
- Box, line, grid and combs remover
- Drop out color
- Dilate
- Erode
- Deskew on border
- Deskew on content

Recognition Engines

Content of preprocessed images is recognized by a powerful set of recognition engines belonging to the following main groups:

- Optical character recognition (OCR) – machine written text
- Intelligent Character Recognition (ICR) – hand written text
- Optical Mark Recognition (OMR) – check boxes
- Bar code, QR code, DataMatrix
- Logo

Papyrus Designer/Classify

The Papyrus Designer/Classify is used to define categories to be classified, the extraction of feature sets and the classification strategy.

The Designer/Classify supports automation of several definition steps. It also enables the easy handling of sample data. The Designer/Classify is fully integrated into Papyrus Objects.

Supported classification steps based on:

- Keywords and its combination
- Layouts
- Advanced text
- Logos

Before you start working with the Designer/Classify, the required document categories have to be defined. With a number of sample documents (a few dozen) for each category the system is trained in order to learn the specific properties of each document class. New categories can also be added easily and conveniently.

The Papyrus Designer/Extract component empowers classification capabilities by letting to define some classification logic even on a later stage – during extraction.

Papyrus Designer/Extract

To discover relevant content the following main possibilities and their mixture are usually used in the extraction definition

- Patterns
- Keywords
- Regions
- Positions
- Anchor points
- Rules
- Layout
- Logo

Basic Definition Process Structure

- Define all document types that occur
- Define the elements of interest on each document type
- Define for each data element:
 - General attributes
 - Related anchors
 - Patterns that apply to the element
 - Condition (rules) that have to be fulfilled
- Test the definitions with various documents

- Compare the results with previous generated “golden files” of intended extract data.

Various display support tools enable you to define new documents fast:

- Display candidates for extraction
- Show regions and coherences
- Convenient tag tools

The Papyrus Capture/Extract module provides a possibility for document reclassification definition or classification definition based on the very deep document structure analysis. This gives a big flexibility for the document design and processing definitions of complex documents.

Data Formats

- TIFF 6.0 or higher (also multi-page)
- JPEG, JPEG 2000, BMP, GIF, PNG
- HTML, CSV, TXT (plain text)
- PDF 1.4 or higher, PDF/A (ISO 19005)
- AFP
- Office formats:
 - DOC/DOCX/DOCM/ODT
 - PPT/PPTX/PPTM/ODP
 - XLS/XLSX/XLSM/ODS

Prerequisites

- Windows 7 or later; Windows Server 2012 or later
- Standard office PC
- Papyrus WebRepository
- Papyrus Capture /Office-In for office formats support
- Papyrus PDF-in for native PDF support

Order Information

Papyrus Designer/Capture
for Windows

Training

ISIS Papyrus offers in-house workshops and standard courses for user training.

Maintenance

The Terms and Conditions for ISIS Papyrus Software Products apply. Free service period is 6 months after installation. A maintenance agreement is offered optionally for updates and product support.

Papyrus Business Designer

User trained Capture, teaching the machine

The screenshot displays the Papyrus Business Designer interface. On the left, a document titled 'Conf_John.tif' is open, showing a confirmation form from St. Francis General Hospital. The form includes fields for patient name (John Avey), date (15.02.2017), and time (04:20 to 04:50 pm). A signature is visible at the bottom right of the document.

In the center, a 'Tester Main View' window shows a table of test results for a 'Medical_confirmation' document type. The table compares a 'Test sample' against a 'Reference run' across five categories: Perfect, Improved, Not different, Worse, and Unknown.

Sample	Perfect	Improved	Not different	Worse	Unknown
Test sample for Conf_John_...	0	0	3	1	0
Test sample for John_Conf_...	0	0	3	1	0
Test sample for Beth_Conf_...	0	0	3	1	0
Test sample for Confirmatio...	0	0	0	0	4

Below the table, a 'Details' section shows the 'Form' tab selected, displaying the document's metadata (ID: UM 066634/16, Name: John Avey, Date: 15.02.2017, Start_time: 04:20). A comparison between 'Golden data' and 'Current run' is shown, indicating a match for the ID and name fields.

On the right, a preview of the document is shown, highlighting the 'CONFIRMATION' section and the patient name 'John Avey'.

Product Description

The **Papyrus Business Designer** is a modern WYSIWYG design tool for easy definitions when digitizing business content with a self-learning user trained capture capability. Document structures and extraction definitions can be created faster and better than ever before.

It specializes in supporting business capture and OCR needs during content and case-management operations, such as extracting/indexing documents in the business case (on the fly, on demand), but also lifting fields from arbitrary documents, again ad-hoc and on the fly. This is possible because the Papyrus Business Designer/Capture does not require IT involvement, it gives all the power into the hands of the actual users of the application.

Any document can be extracted on the fly, regardless whether a document template has been created for it or not. Extraction definitions created by one user can be shared with other users or entire departments.

To lift a field from a document, the business user simply drags the text/area that is to be extracted from the document onto a pane. For its OCR/ICR/Classification capabilities the Papyrus Business Designer relies on the IDEX-Engine, also used in the Papyrus Recognition Server, which is a prerequisite for the deployment of Papyrus Business Designer.

Contact

Europe

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

America

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

Asia Pacific

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

Email & Web

info@isis-papyrus.com
www.isis-papyrus.com

© 2020, ISIS Papyrus,
replaces all previous
documentation.

Key Benefits

- User trained capture capability
- Increased efficiency for content digitizing: document types and extraction definition can be created faster and better than ever before.
- Easy and clear application for business users supported by user-trained technologies to create an extraction definition
- Excellent tool for both business and IT allowing a perfect collaboration
- Optimal collaboration and assuring the best results by utilizing Papyrus change management
- Recognition quality assurance supported by automatic evaluation of the extraction results for supervisor decision making

Functions

- Document types and extraction definitions for business users
- User trained extraction with sample documents
- Change and release management of document definitions
- Recognition quality evaluation

Document Extraction Definitions

- Upload document samples
- Create new document types
- Create new data fields and UI form elements: specify element type, region, keywords, etc.
- Create new table data fields: specify column element type, table header keywords etc.
- Extract and paste document content into an existing form or table field by drag and drop
- User trained document definitions fully compatible with Papyrus Capture Designer allowing to refine certain field definitions by extraction development experts for best performances
- Rearrange or delete fields from form, delete table columns

Training

- Create implicit training data during creation of new fields
- Train extraction definitions with training data collected during fields manipulations (creation or update)
- Refine extraction definitions by deleting training data and retraining

Classification

- Classify new document samples
- Correct document sample categories

Change Management

- Manage full life cycle of document structure and extraction definition
- Define change management process: private and group development
- Deactivate or restore document and extraction definitions to last productive version
- Delete not productively used document definitions

Recognition quality assurance

- Recognition results representation for the trained document samples
- Automatic compare with the reference data from the previous productive version of the extraction definition
- Representation of the extraction and comparison results in the optimal way to the supervisor

Integration

Papyrus Business Designer/Capture is natively integrated within Papyrus Adaptive Case Management (ACM) and is fully supported by it.

IT-experts can do extensions of the created extraction definition also with the Papyrus Designer Capture in case complex extraction logic is required.

Formats

- TIFF 6.0 or higher (also multi-page)
- JPEG
- BMP
- GIF
- PNG
- PDF 1.4 or higher
- PDF/A (ISO 19005)
- Office formats:
 - DOC/DOCX/DOCM/ODT
 - PPT/PPTX/PPTM/ODP
 - XLS/XLSX/XLSM/ODS

Prerequisites

- Windows 7 or later, Windows Server 2012 or later
- Standard office PC
- Minimum display resolution of 1280x1024
- Papyrus WebRepository
- Papyrus Recognition Server
- Papyrus Capture /Office-In for office formats support
- Papyrus PDF-in for native PDF support
- Anonymized document samples for training (GPDR conform)

Order Information

Papyrus Business Designer
for Windows

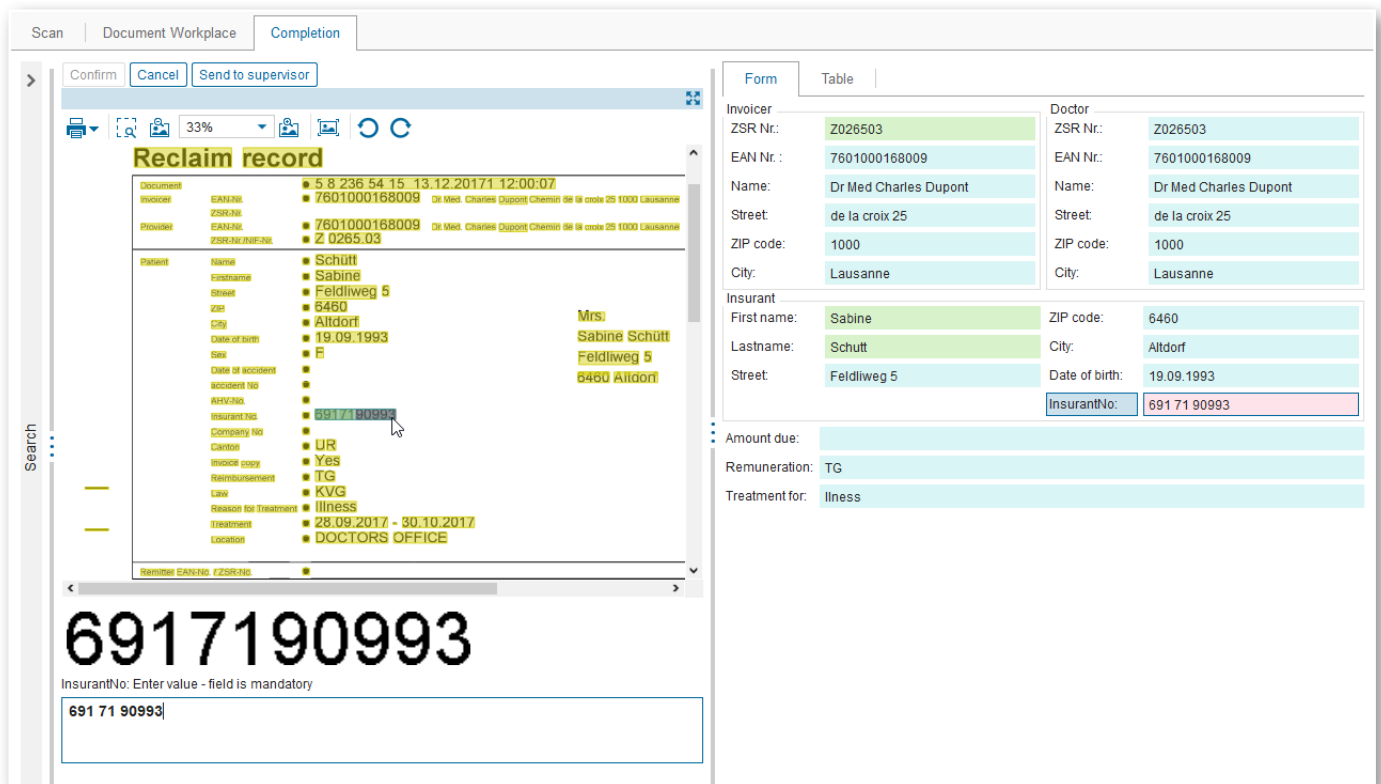
Training

ISIS Papyrus offers in-house workshops and standard courses for user training.

Maintenance

The Terms and Conditions for ISIS Papyrus Software Products apply. Free service period is 6 months after installation. A maintenance agreement is offered optionally for updates and product support.

Papyrus Client/Capture



Product Description

Papyrus Client/Capture is the Papyrus EYE Widgets based end-user interface for all kinds of user interactions within a document capture system. It enables the completion and correction of classification and extraction results with all necessary display and validation functions:

- Uniform handling of all incoming documents
- Interactive check and correction of extraction results
- Immediate verification of corrected data
- Flexible image display
- Information messages indicating recognition errors and rule violations
- Administrator and operator views

Papyrus Client/Capture can handle expanded requirements for manual input processing such as rearranging the sequence of sheets or exchanging pages between documents.

Document Verification and Correction Interface for Capture Operator Powered by Validation

Contact

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

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

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

Email & Web
info@isis-papyrus.com
www.isis-papyrus.com

© 2020, ISIS Papyrus,
replaces all previous
documentation.

Benefits

- Easy access to all documents via Papyrus Desktop/EYE Widgets or a web-browser connected to a **Papyrus WebPortal**
- One user interface for all types of documents and all tasks
- Workplaces with pre-configured views for capture operators and administrators
- Customizable to your individual needs without programming
- Ergonomic screen designs

Core Features

- Display, view and edit all types of processed inbound documents
- Handling of virtually any document format – images (TIFF, JPEG, BMP, GIF, PNG), HTML, CSV, TXT, native PDF and PDF/A, AFP as well as all common office formats.
- Fine-tuned search function for Capture documents – search for document type, status, priority and name
- Thumbnails Viewer with integrated document manipulation functions
- Image Viewer with convenient zooming and printing functions
- Image Detail Viewer and color highlighting for areas of special interest
- Variable color highlighting for data validation
- Visual marking of rejected characters
- Freely configurable masks for data display
- Online validation with assistant for substitution lookup

Capture Framework Integration

For rapid start-up the **Papyrus Client/Capture** includes so-called capture workplaces with pre-configured views for capture operators and administrators, which provide a complete set of functionalities required for a successful capture project.

- The **Capture Workplace** provides views for scanning, document management and document completion
- The **Capture Administration Workplace** offers views for administrative tasks such as queue and adapter control, process monitoring and keeping track of tasks in the workflow

Document Workplace View

The **Papyrus Client/Capture** provides the business user with a Document Workplace view which enables the business user to easily manage and manipulate all processed inbound documents:

- classify documents manually
- split and combine documents
- rotate, move and delete pages
- move images to a clipboard
- scan/import additional pages

Prerequisites

- Windows 7 or later, Windows Server 2012 or later
- Standard office PC
- Papyrus WebRepository
- Papyrus Capture /Office-In for office formats support

Order Information

Papyrus Client/Capture 7
for Windows

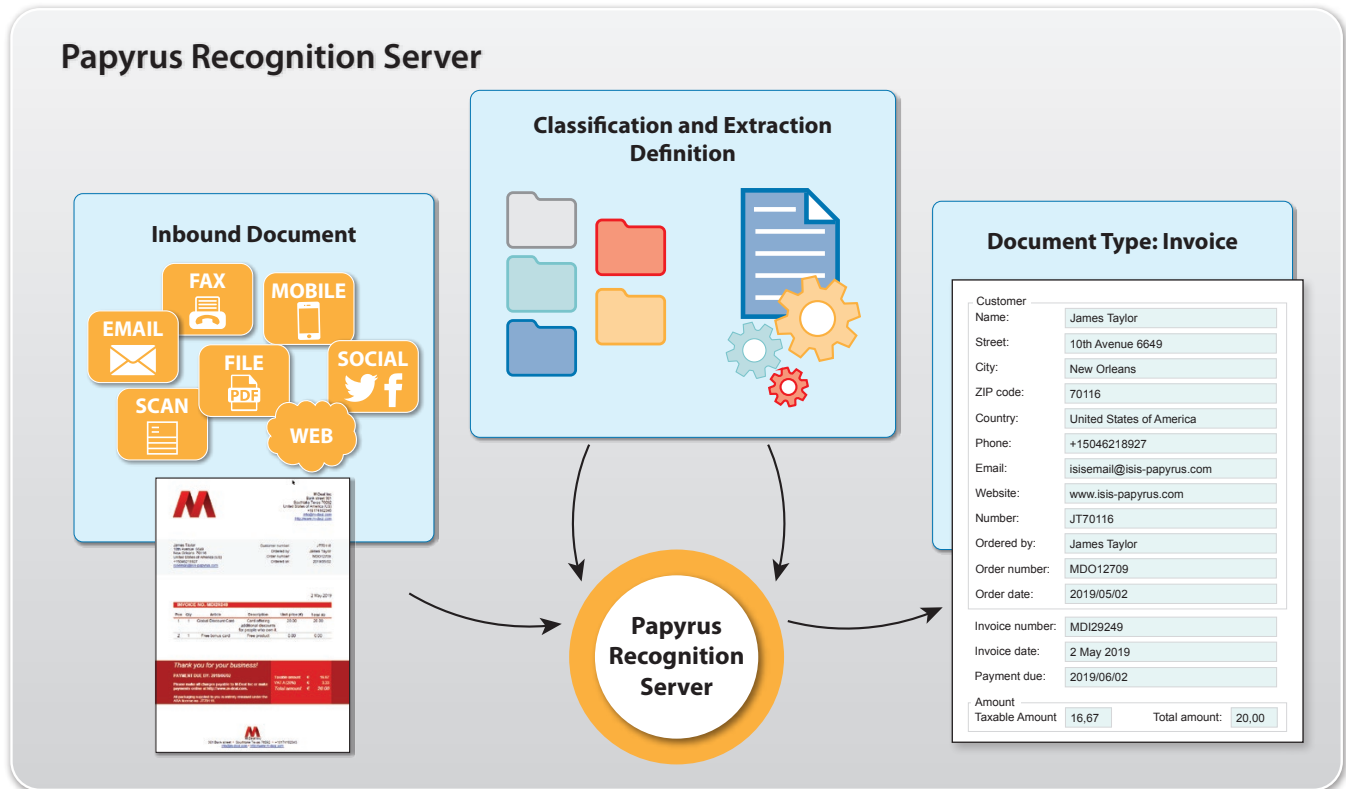
Training

ISIS Papyrus offers in-house workshops and standard courses for user training.

Maintenance

The Terms and Conditions for ISIS Papyrus Software Products apply. Free service period is 6 months after installation. A maintenance agreement is offered optionally for updates and hotline.

Papyrus Recognition Server



Product Description

Papyrus Recognition Server is a powerful tool for automated high volume business content digitizing what is key for tight SLAs (service level agreements) in the modern business world.

The Recognition Server offers a broad range of possibilities for application in the field of automated sorting and distribution of electronic documents, fax and paper mail. For data extraction, the Recognition Server can identify unstructured and structured documents with great reliability. Various types of content: whether machine or hand written text, check boxes or diverse bar codes, QR codes, DataMatrixes - can be recognized and identified as business related content by means of powerful document analysis using pattern recognition, keywords, relative positions, voting algorithms, layout analysis and rules.

Papyrus Recognition Server applies classification and extraction definitions created with the Papyrus Designer Capture or Papyrus Business Designer /Capture to acquire the required business content from the inbound business mails.

Key Benefits

- Automated high volume content digitizing to keep business SLAs
- Powerful classification and extraction of business relevant content out of high volume of inbound data
- Excellent results for both FixForm and FreeForm® documents
- Support of documents from all common file formats and inbound channels such as scan, e-mail, mobile/photos, web and social media.
- Support of diverse recognition engines for different kinds of content, powerful preprocessing and extraction defined with Papyrus Designer Capture and Papyrus Business Designer.

Contact

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

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

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

Email & Web
info@isis-papyrus.com
www.isis-papyrus.com

© 2020, ISIS Papyrus,
replaces all previous
documentation.

Features

Document Classification

The purpose of the Papyrus Recognition Server for document classification is to enable an automated process of judging incoming documents according to selected criteria, sorting them into freely definable categories – thereby making information actually accessible for the company – and forwarding them accurately to those who are in charge of the issue.

Documents which could not be assigned to a class unequivocally are separated and presented to the administrator or clerk for further assessment.

Data Extraction

The Papyrus Recognition Server extracts and reads all necessary field data from the identified document class. This process can include unstructured layouts and structured forms with keywords at predefined positions.

Image Preprocessing

According to the applied classification and extraction definitions the following image pre-processing can be used:

- Dirt removal, Despeckle
- Punch hole removal
- Automatic rotation
- Binarize
- Convert color to gray
- Box, line, grid and combs remover
- Drop out color
- Dilate
- Erode
- Deskew on border
- Deskew on content

Recognition engines

Content of pre-processed images can be recognized by a powerful set of recognition engines belonging to the following main groups:

- Optical character recognition (OCR) – machine written text
- Intelligent Character Recognition (ICR) – hand written text
- Optical Mark Recognition (OMR) – check boxes
- Bar code, QR code, DataMatrix
- Logo
- Signature

Document structure analysis

To discover relevant content the following main possibilities and their mixture can be applied during the extraction process based on the prepared extraction definition:

- Patterns
- Keywords
- Regions
- Positions
- Anchor points
- Rules
- Layout
- Logo

Data formats

The Papyrus Recognition Server can process the following data formats:

- TIFF 6.0 or higher (also multi-page)
- JPEG and JPEG 2000
- BMP
- GIF
- PNG
- HTML
- CSV
- TXT (plain text)
- PDF 1.4 or higher
- PDF/A (ISO 19005)
- AFP
- Office formats:
 - DOC/DOCX/DOCM/ODT
 - PPT/PPTX/PPTM/ODP
 - XLS/XLSX/XLSM/ODS

Prerequisites

- Windows 7 or later; Windows Server 2012 or later
- Linux (SLES 11 or later, RHEL 5 or later)
- Papyrus WebRepository
- Papyrus Capture /Office-In for office formats support
- Papyrus PDF-in for native PDF support
- Papyrus Designer Capture or Papyrus Business Designer

Order Information

Papyrus Recognition Server for Windows and Linux

Training

ISIS Papyrus offers in-house workshops and standard courses for user training.

Maintenance

The Terms and Conditions for ISIS Papyrus Software Products apply. Free service period is 6 months after installation. A maintenance agreement is offered optionally for updates and hotline.