

# Papyrus Capture - Intelligent Paper Conversion

ISIS Papyrus.  
Business Document Automation and Management™.

## Getting airborne cost effectively with Papyrus Capture

### Application report Lufthansa AirPlus

With their smart company card and employee card system, Lufthansa offers a modern and transparent management of corporate travel expenses. The company card system offers travel services through the individual's travel agency while billing is done by AirPlus. This way detailed account information with additional data like cost centers and insurance are integrated into the statement. The system is designed to reduce travel related expenses and to optimize the overall process of business travel, which proved to be highly successful. The number of cards issued in 1998 increased by 27% to approximately 386,000.

The success of the company card program and the related flood of incoming application forms made higher data transfer rates and a modernization of the data capturing process a matter of utmost urgency. Selection criteria were clear: The system would have to provide smooth upward scalability and a high degree of flexibility regarding the definition of new forms.



From the early planning stage on, there was a close cooperation with Lufthansa regarding a 'reader friendly' design of the application forms. Detailed engineering specifications provided a solid base for the system solution. The system is completely integrated in Lufthansa's standard computer environment: NT servers run the data and document management as well as the recognition process. The scanning of forms and the post processing of the captured data is done on standard NT workstations in the assigned company division.

Since most of the forms are filled out by hand, a high performance in hand print recognition is extremely important. Improx' reliable ICR engine is supported by reference tables of German city and street names. The matching procedure uses 'fuzzy logic' so that even incompletely read names are recognized successfully. This technology reduces the manual capture and correction effort considerably.

Papyrus Capture uses cutting edge technology which provides a solid base for the capture of tens of thousands of data sets per day. The system expands easily with the growing success of Lufthansa AirPlus and is capable to meet the challenges of supporting new card products with reliable data capture.

## More Customers

### Deutsche Post AG

Papyrus Capture is used to automatically record orders for package transportation since May 1999. An average of 35,000 to 40,000 documents daily are scanned with 33 Kodak DS 9500 scanners. After automatic extraction, the data is passed on to an SAP system, an IXOS archive as well as to the tracking & tracing system of the German postal service.

### RSC Raiffeisen Data Service Center

RSC is using Papyrus Capture since April 1999 to record more than 20,000 deposit slips, debit slips and checks daily. After being scanned in special A6 item scanners, the documents are processed through a complex workflow which includes a three stage post processing. Then the data is exported to the Raiffeisen accounting system in DTA format.

### GbD - Institute for Paperless Data Processing

This service company records up to 200,000 flight coupons daily for various airlines. The IBML Image Trac high performance scanner provides color images to be used by Papyrus Capture to extract the relevant data automatically and pass them on to the booking system of the appropriate airline.

### DRK Blood Bank Baden Württemberg

Blood donor data

### GTG Bank Services

Deposit slips, debit slips, checks and travelers checks

### EMD - Electronic and Microfilm Documentation Systems

Keyword catalog of the Austrian National Library

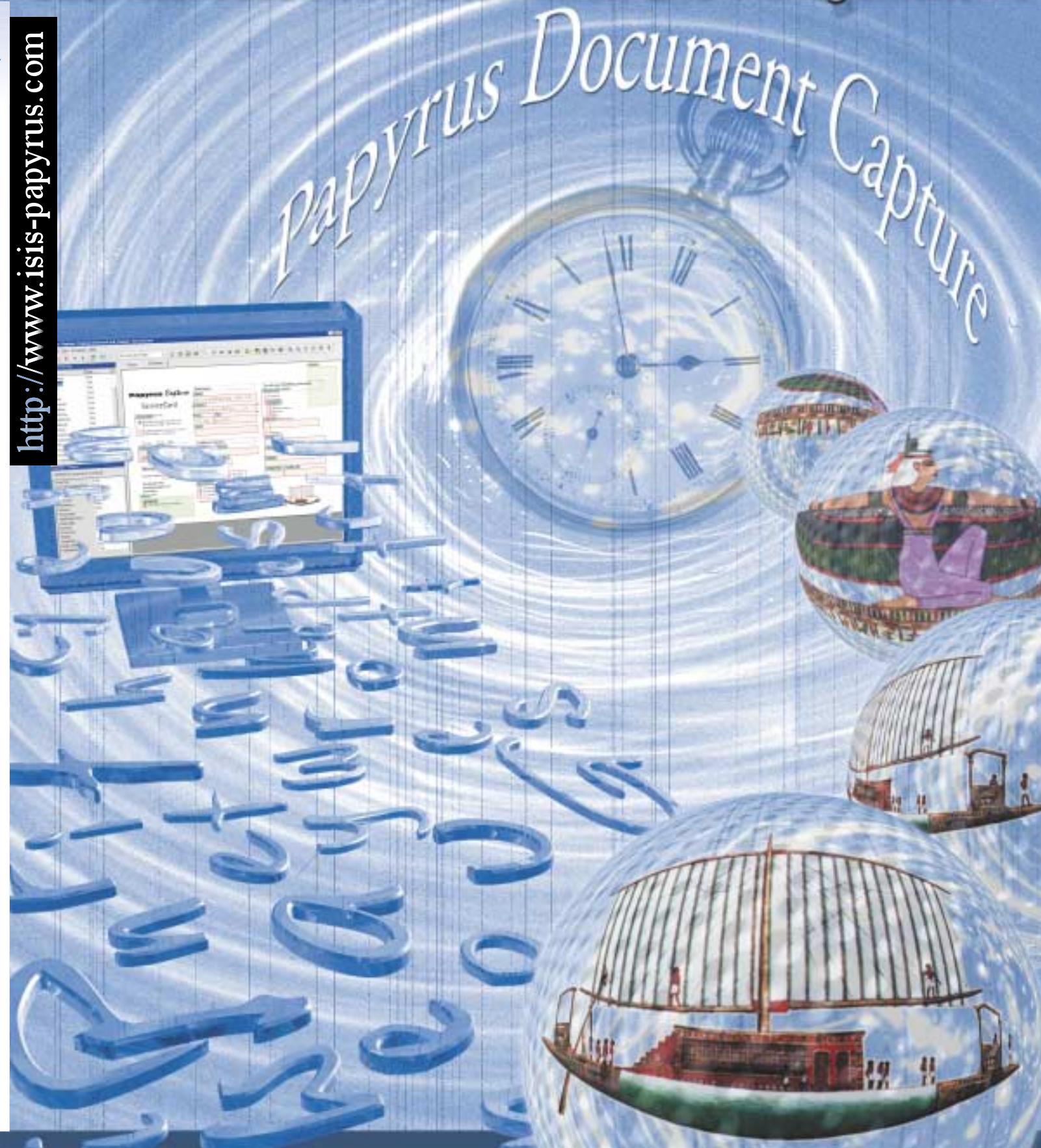
### Canon

Counter status cards

### DWS German Institute for Security Savings

German tax relieve order

<http://www.isis-papyrus.com>



Evolution did not stop with walking upright.

## The Concept

**Papyrus Capture** is a high performance system for automatic recognition and capture of business documents which translates printed documents into electronic data. It is based on the extremely successful BlueWings software by Improx. The system recognizes and verifies the content and provides the retrieved information to other applications for further processing. The individual steps - scanning or fax import, classification of the business documents, recognition, verification and data export have been automated in **Papyrus Capture** to the highest possible degree.

### Advantages

The capture of large quantities of document print such as order forms, application forms, invoices, money transfer forms, questionnaires and other business documents requires considerable expenses in terms of time and human resources. This major expense of any business operation gets efficiently cut by **Papyrus Capture** which often amortizes after between 6 and 12 months time. This fact is based on the following advantages:

- ⊕ Speed - substantial acceleration of the document transformation process; data becomes available faster
- ⊕ Cost - drastic reduction of expenses for data capture and data entry
- ⊕ Quality - improvement of data quality through automatic plausibility checks
- ⊕ Personnel - employees are relieved from performing time-consuming routine tasks. They can focus on using relevant data instead of generating them by hand
- ⊕ Versatility - access to data from archives and work flow systems
- ⊕ Additional benefits - a higher percentage of a company's stored data is getting handled and put to use

## The Architecture

Document processing with **Papyrus Capture** follows the basic steps of scanning or data/fax import, recognition, completion and export. These steps are carried out by the respective **Papyrus Capture** data processing modules Scan, Import, Reco, Verify and Export, which together form the capture workflow. All data processing modules have direct access to the **Papyrus Capture Base Services**. Combined with the scripting interface, this allows for a largely free configuration of the individual modules. For instance, documents can be recognized as a whole or in part already at the time of scanning, or recognition services can be evoked during verification.

In conjunction with an almost free choice in combining individual modules, the implementation of very complex workflows becomes possible, which is a distinguishing feature of **Papyrus Capture** compared to other, conventional document capture systems.

### Application Design

The designer is a service for the creation and maintenance of **Papyrus Capture** applications (projects). During the design of an application specific requirements are listed, individual parameters are specified and the following definitions are introduced:

- ⊕ Document classes with field descriptions, classification rules, extraction rules, plausibility rules and dialogues
- ⊕ Processing modules and capture workflow
- ⊕ Parameter objects and clients
- ⊕ Capture assistants
- ⊕ Statistic reports for continuous surveillance and optimizing of the system

## System Administration

The System Administration observes the actual field performance of **Papyrus Capture** and, should the need arise, introduces corrections and alterations. This includes the re-adjustment of various processing parameters (such as settings, assistants and scripts) as well as changing the status of the monitored document.



### The Work Area

The user acts exclusively in the work area, which gives him only limited access, according to his user rights that are preset by the administrator.

### The Scan Module

The Scan Module drives the scanner and stores the scanned document images in single or batch mode in the **Papyrus Capture** database.

If needed, the Scan Module can also perform structuring tasks for the document workflow. Information necessary for providing this service is called up via the recognition services - which can be accessed from any module as part of the **Papyrus Capture Base Services** - from the image content of the document image. This way documents can be preclassified during the scanning procedure.

### The Import Module

Imports document images from an external source and saves them one by one or in batches in the **Papyrus Capture** document database. If necessary, the Import Module can also perform recognition and structuring tasks on incoming documents.

### The Recognition Module

Is used to recognize scanned document images. The Recognition Module provides a data base validation during the recognition process, which allows for vastly improved effectiveness.

### The Verify Module

Provides manual post processing and inspection of the documents which could not be uncontestedly read or validated by **Papyrus Capture**. A screen message notifies the user of characters, fields and documents which do not comply with the plausibility rules laid out in the document definition.

### The Export Module

Formats recognized and validated document data and transfers it with or without the original image to the desired place. Any number of export modules can be used (cascaded) to enable multiple export. Export modules can also be configured for end of work day execution.

## FreeForm

**FreeForm®** is an option to **Papyrus Capture** which enables extensively automated processing of unstructured documents. It is based on a two-phase operation: Classification and extraction. The classification stage identifies the type of the document (e. g. delivery notice, invoice, application form, etc.) according to the read document content. In this stage various partly self-learning procedures are employed. The following stage of extraction recognizes the relevant defined fields. In contrast to conventional automatic forms processing, the exact position of individual fields does not have to be known; it is automatically determined according to predefined rules.

**FreeForm®** features a convenient designer for quick and effective definition of relevant rules as well as to control and monitor the self learning process.



### Advantages

- ⊕ Processing of common business documents - not just forms
- ⊕ Recognition technology does not rely on predefined reading zones
- ⊕ Simple definition of document classes

### Applications

- ⊕ Incoming mail scanning with automatic sorting and distribution
- ⊕ Recording of applications, invoices, contracts, etc.
- ⊕ Automatic document classification according to keywords and phrases

### Functions

- ⊕ Automatic data extraction by interpretation of the full text within a self learned set of rules
- ⊕ The following data types are found in the document and semantically assigned and checked: Numbers, text, date address, tables and other user defined custom types