

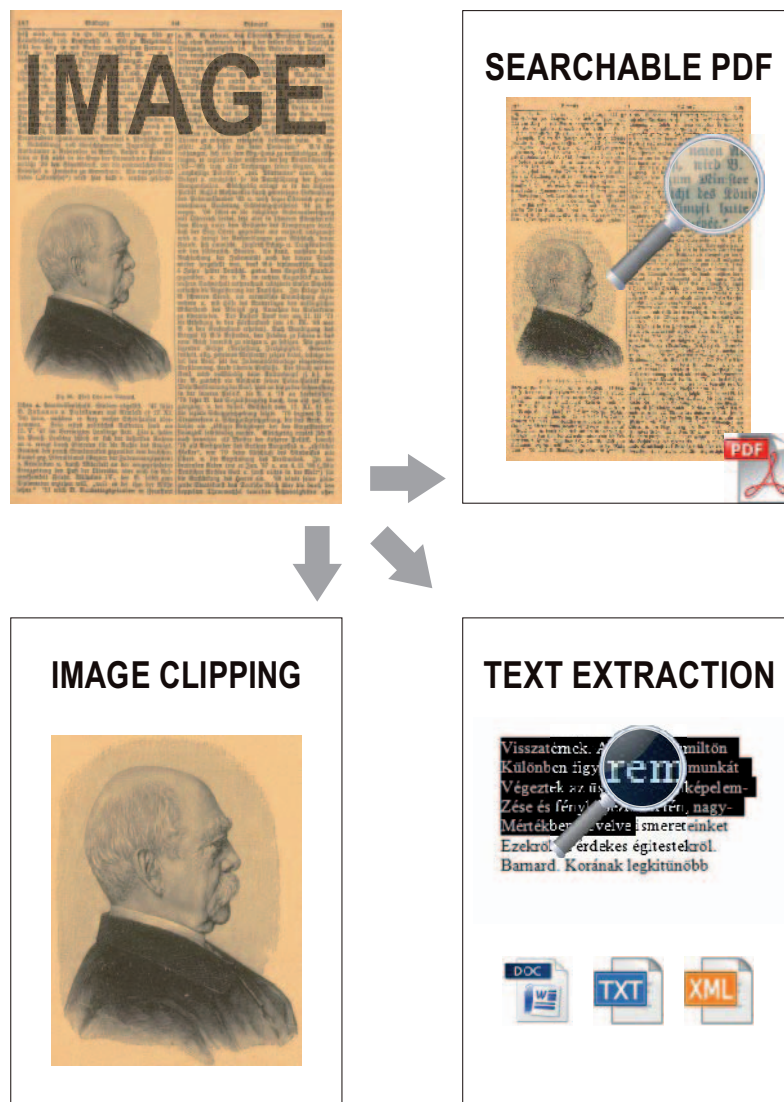
TREVENTUS

# OCRSOLUTIONS

Software for automated text recognition

## Advantages

- Highly automated - no operator interaction needed
- Direct integration into digitization workflows
- Various OCR packages for individual project sizes and requirements
- Entire solution from one supplier



TREVENTUS

# OCR Solutions by TREVENTUS

## General product overview

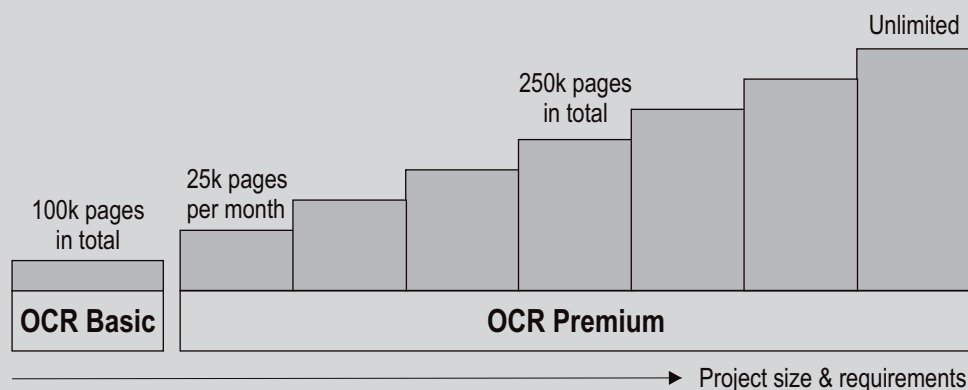
Depending on the requirements of the digitizing projects different OCR - Treventus packages are available:

### I. OCR - Treventus Basic

Software solution for small and medium projects with simple OCR requirements based on "ABBYY FineReader Corporate".

### II. OCR - Treventus Premium

Software solution for medium to large projects with complex OCR requirements based on "ABBYY Recognition Server".

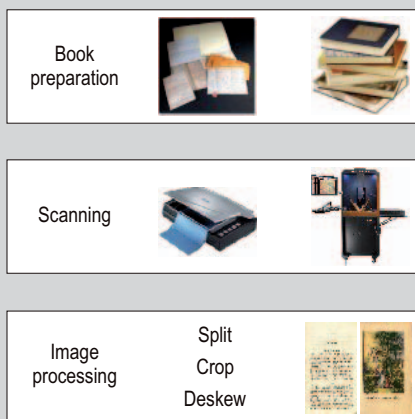


## Benefits

- Simple user interface for convenient use with direct integration into digitization workflows (directly implemented in ScanGate™)
- More than 190 supported languages
- Wide range of output formats
- Automated processing - 24/7
- Installation support due to individual needs

# Workflow example including OCR

## Tasks before OCR



## Automated OCR - Treventus Basic / Premium

### SEARCHABLE PDF



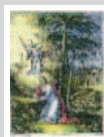
### TEXT EXTRACTION



### OPTIMIZED PDF



### IMAGE CLIPPING



### OCR correction<sup>1</sup>



manual task

<sup>1</sup> OCR - Treventus Premium

## Final document storage

Archive

Backup

**NAINUWA**<sup>TM</sup>  
*Digital Library*

# OCR - Treventus Basic vs. Premium

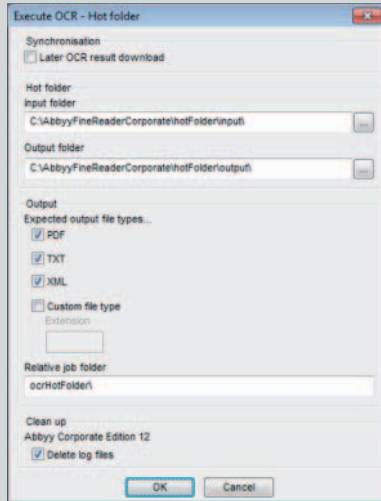
## Comparison chart

Functions / Features	OCR - Treventus Basic	OCR - Treventus Premium
<b>General</b>		
Automated OCR processing via Hot-folder	●	●
ABBYY FineReader Corporate included	●	–
ABBYY Recognition Server included	–	●
Availability and access to OCR-software updates	–	●
<b>Interfaces</b>		
OCR - Treventus Basic interface in ScanGate™	●	–
OCR - Treventus Premium interface in ScanGate™	–	●
OCR settings customizable directly in ScanGate™	–	●
Remote Admin-console	–	●
<b>Performance</b>		
Scalability (a): various licence options	–	●
Scalability (b): performance upgrade options <sup>1</sup>	–	●
24/7 fail-proof large volume processing	–	●
<b>Output</b>		
Output formats (a): pdf, txt, xml	●	●
Output formats (b): pdf/a, doc and epub	–	●
Output formats (c): Nainuwa™ - OCR-files	●	●
File naming customizable (based on meta data)	–	●
Creation of multi-page files of (a): pdf, txt and xml	●	●
Creation of multi-page files of (b): tif, pdf/a, doc, rtf	–	●
Creation of multi-page files of (b): html, epub, xml alto	–	●
<b>Extra functions</b>		
OCR - verification possibility	–	●
Support of XML-tickets	–	●
OCR - exception handling	–	●
Text extraction	●	●
Image clipping	–	●
<b>Historic Fonts</b>		
Gothic font (Fraktur) OCR <sup>1</sup>	–	●

<sup>1</sup>Optional acquirable

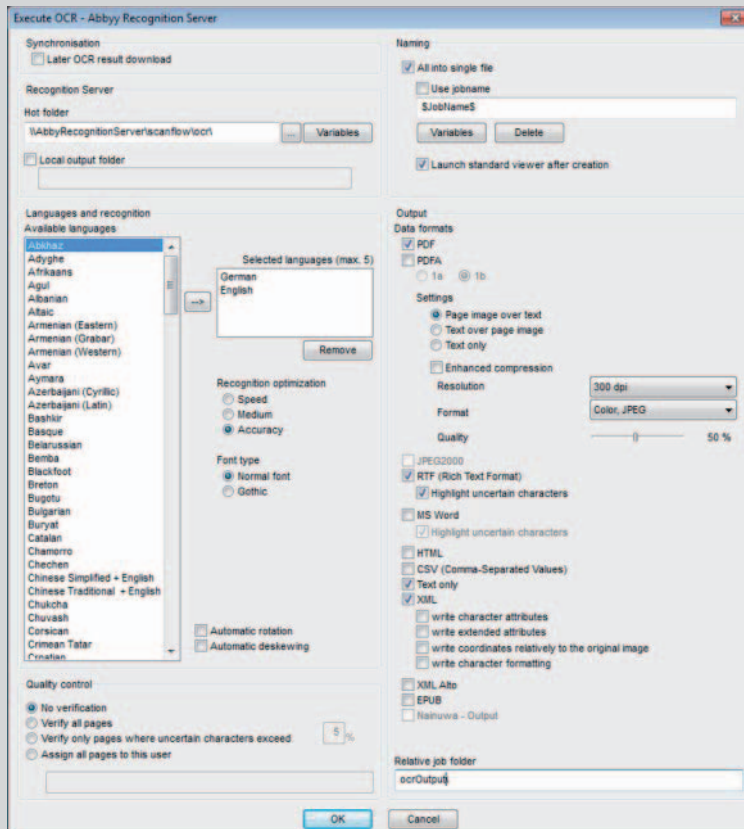
# OCR Interface module - Basic vs. Premium

## OCR - Treventus Basic user interface in ScanGate™



Only the basic settings can be chosen directly from ScanGate™ / ScanFlow™

## OCR - Treventus Premium user interface in ScanGate™



All OCR settings can be chosen directly from ScanGate™ / ScanFlow™

# OCR - Treventus specifications

## Supported recognition languages

### + 43 main languages with dictionary support:

Arabic<sup>1</sup> (Saudi Arabia), Armenian (Eastern), Armenian (Grabar), Armenian (Western), Azeri (Latin), Bashkir, Bulgarian, Catalan, Croatian, Czech, Danish, Dutch, Dutch (Belgian), English, Estonian, Finnish, French, German, German (new spelling), Greek, Hebrew, Hungarian, Indonesian, Italian, Latvian, Lithuanian, Norwegian, Norwegian (Bokmal), Norwegian (Nynorsk), Polish, Portuguese, Portuguese (Brazilian), Romanian, Russian, Slovak, Slovenian, Spanish, Swedish, Tatar, Thai, Turkish, Ukrainian, Vietnamese<sup>1</sup>;

### + 133 additional languages without dictionary support:

Abkhaz, Adyghe, Afrikaans, Agul, Albanian, Altai, Avar, Aymara, Azerbaijani (Cyrillic), Basque, Belarusian, Bemba, Blackfoot, Breton, Bugotu, Buryat, Cebuano, Chamorro, Chechen, Chukchee, Chuvash, Corsican, Crimean Tatar, Crow, Dargwa, Dungan, Eskimo (Cyrillic), Eskimo (Latin), Even, Evenki, Faroese, Fijian, Frisian, Friulian, Gagauz, Galician, Ganda, German (Luxembourg), Guarani, Hani, Hausa, Hawaiian, Icelandic, Indonesian, Ingush, Irish, Jingpo, Kabardian, Kalmyk, Karachay-balkar, Karakalpak, Kasub, Kawa, Kazakh, Khakass, Khanty, Kikuyu, Kirghiz, Kongo, Koryak, Kpelle, Kumyk, Kurdish, Lak, Latin, Lezgi, Luba, Macedonian, Malagasy, Malay (Malaysian), Malinke, Maltese, Mansi, Maori, Mari, Maya, Miao, Minangkabau, Mohawk, Moldavian, Mongol, Mordvin, Nahuatl, Nenets, Nivkh, Nogay, Nyanja, Ojibway, Ossetian, Papiamentu, Provençal, Quechua, Rhaeto-Romanic, Romany, Rundi, Russian (Old Spelling), Rwanda, Sami (Lappish), Samoan, Scottish Gaelic, Selkup, Serbian (Cyrillic), Serbian (Latin), Shona, Sioux (Dakota), Somali, Sorbian, Sotho, Sunda, Swahili, Swazi, Tabasaran, Tagalog, Tahitian, Tajik, Tok Pisin, Tongan, Tswana, Tun, Turkmen, Tuvian, Udmurt, Uigur (Cyrillic), Uigur (Latin), Uzbek (Cyrillic), Uzbek (Latin), Welsh, Wolof, Xhosa, Yakut, Yiddish, Zapotec, and Zulu;

### + 5 East Asian languages:

Chinese<sup>1</sup> (Traditional, Simplified), Japanese<sup>1</sup>, Korean<sup>1</sup> and Hangeul<sup>1</sup> (Korean);

### + Gothic print<sup>1</sup> (Fraktur):

for English, French, German, Italian, Spanish, Latvian;

### + 4 artificial languages:

Esperanto, Ido, Interlingua, and Occidental;

### + 6 programming languages:

Basic, C/C++, COBOL, Fortran, Java, and Pascal;

<sup>1</sup>Optional acquirable

## General system requirements

**Processor:** Intel® QuadCore, 2.0 GHz or above

**Operating system:** Microsoft® Windows® 7 / 8

**Memory:** 4 GB RAM

**Hard Disk Space:** 1 TB or above (recommendation)

**Microsoft .NET Framework:** 2.5 or above

# OCR - Treventus product overview

Contact us for your individual offer!

Product name	OCR contingent
OCR - Treventus Premium Unlimited	Unlimited pages <sup>1</sup>
OCR - Treventus Premium 1000k	1,000,000 pages total <sup>2</sup>
OCR - Treventus Premium 500k	500,000 pages total <sup>2</sup>
OCR - Treventus Premium 250k	250,000 pages total <sup>2</sup>
OCR - Treventus Premium 100k / month	100,000 pages / month <sup>3</sup>
OCR - Treventus Premium 50k / month	50,000 pages / month <sup>3</sup>
OCR - Treventus Premium 25k / month	25,000 pages / month <sup>3</sup>
OCR - Treventus Basic 100k	100,000 pages total <sup>4</sup>
OCR - Premium Interface Module <sup>5</sup>	n.a. <sup>5</sup>

<sup>1</sup> No page limitation for the OCR - processing with 1 Dual Core licence.

<sup>2</sup> Total OCR contingent in A4-format (not time limited) without core limitation.

<sup>3</sup> Monthly OCR contingent in A4-format (not time limited) with 1 Dual Core licence.

<sup>4</sup> Total OCR contingent in A4-format (not time limited).

<sup>5</sup> This module is meant for customers which already have an "ABBYY Recognition Server 3.5 or above".  
The OCR contingent is therefore limited to the given licence of the customer.

This module package moreover includes the configuration of the interface and a remote installation support.

*Annotation: Hardware and ScanGate™ or ScanFlow™ licences are not included within the listed packages.*

## Optionals for OCR - Treventus Premium

### I. Gothic font packages

- + ) 250,000 pages
- + ) 500,000 pages
- + ) 750,000 pages
- + ) 1,000,000 pages

### II. Language packages

- + ) Arabic
- + ) Chinese, Korean & Japanese
- + ) Vietnamese

### III. Use of additional cores

- + ) 1x Dual core
- + ) 1x Quad core
- + ) More cores on request




# PRODUCT FAMILY

**SCANROBOT® 2.0 MDS**  
MASS DIGITIZATION SYSTEM

Automatic scanning  
fast and gentle up to 2.500 pages / hour

Semi-automatic scanning  
for most gentle scanning of fragile books

Manual scanning  
covers, spine, fold outs...



UNIQUE PATENTED PRISM TECHNOLOGY

TREVENTUS

**SCANGATE™**  
Capturing, Processing & Management Software

Advantages


- Easy and intuitive
- Web-based interface
- Applicable for different scanner hardware
- Real-time preview for fast quality control
- Comprehensive processing functions



... with integrated workflow module

TREVENTUS

**SCANFLOW™ Professional**  
Workflow & Management Software



Advantages

- Increasing the throughput & robot quality
- Efficient organization of staff, scanning features, data and books
- Minimization of human interaction and errors
- Fast quality control functions

... let the software work


TREVENTUS

**NAINUWA™**  
*open your Mind*

Dynamic Digital Collection Viewer

Advantages

- innovative image stitching
- flexible search and sort functions
- fast and efficient image loading
- responsive scanning
- ergonomic view handling
- browser-based application



TREVENTUS

**CONSULTING SERVICE**  
*support at any stage*

TREVENTUS, your...

- One stop shop for your digitization project
- Rapid needs analysis search for your needs
- Rapid needs analysis search for your needs
- Experienced and professional project management
- State of the art technology

**BUILD YOUR DIGITAL LIBRARY!**

Digital Library System	Storage
Analysis	Documentation
OCR	Processing
Quality check	Scanner hardware
Backup	Workflow management
Book logistic	Scanning
Plant	Organize

... ask us for advice

TREVENTUS

**TREVENTUS OCR SOLUTIONS**  
Software for automated text recognition

Advantages

- Highly accurate - no operator interaction needed
- Direct integration into digitization workflow
- Various OCR packages for individual project sizes and requirements
- Extensive customer service package



TREVENTUS



## Awards for TREVENTUS

Winner of the European ICT Grand Prize  
Innovation prize of the Theodor Kery foundation  
1<sup>st</sup> place in Genius Innovation Award

TREVENTUS Mechatronics GmbH  
Siebenbrunnengasse 17 / Top 2  
1050 Vienna - AUSTRIA

Tel: +43 1 890 35 10-02  
Fax: +43 1 890 35 10-15

E-Mail: sales@treventus.com  
www.treventus.com