

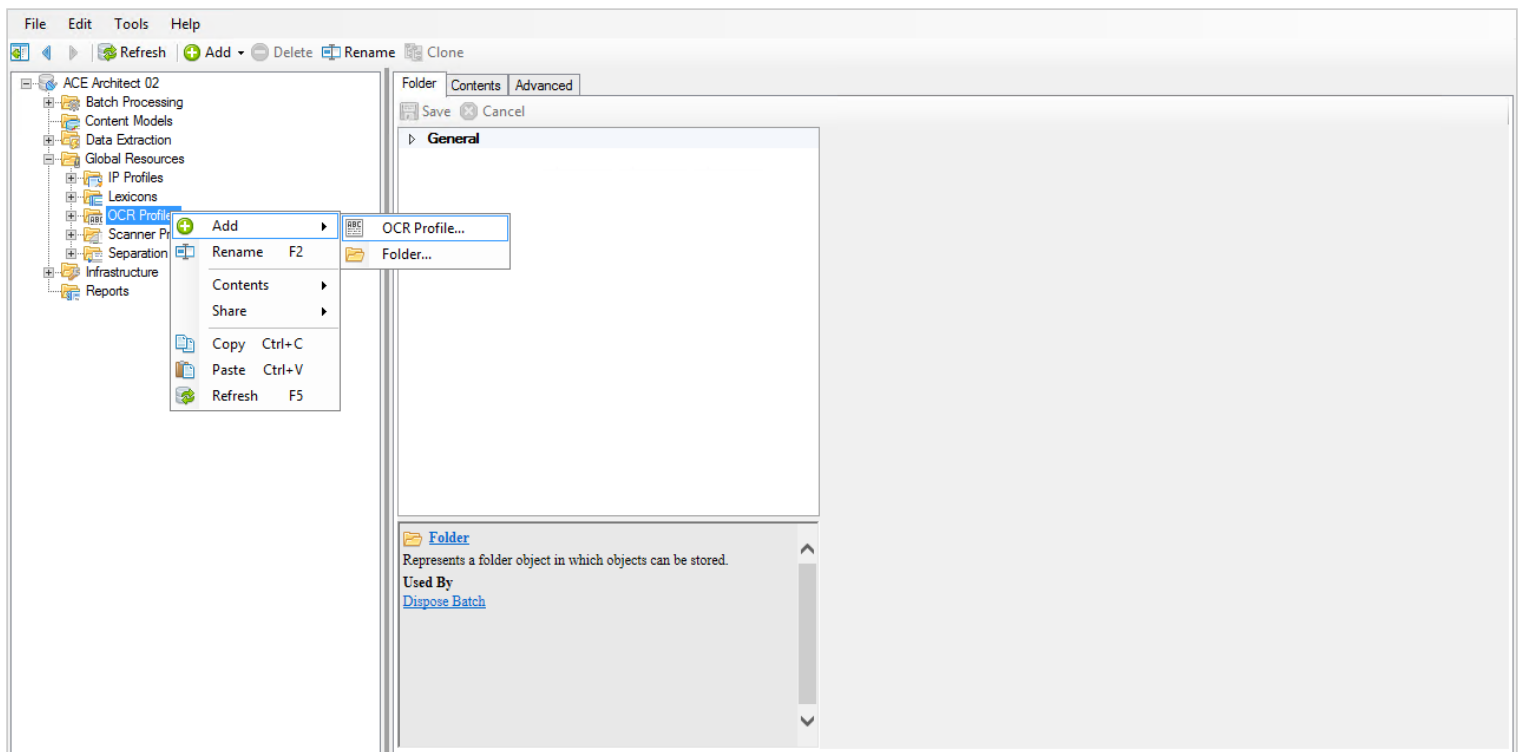
Condition

OCR

Creating an OCR Profile

> Global Resources / **OCR Profiles**

1. Right click and select **Add > OCR Profile...** .
2. Name it **Invoices OCR** .



FileEditToolsHelp

RefreshAddDeleteRenameClone

ACE Architect 02

- Batch Processing
- Content Models
- Data Extraction
- Global Resources
- IP Profiles
- Lexicons
- OCR Profiles
 - Downloads
 - Invoices OCR
- Scanner Profiles
- Separation Profiles
- Infrastructure
- Reports

OCR ProfileOCR TestingAdvanced

SaveCancel

General

OCR EngineTranssym OCR 4

IP Profile

Region of Interest(0,0):(0,0)

Bound Region ProcessingDisabled

Description

Synthesis Options

SynthesisEnabled

Font PitchAuto

Maximum Variance12%

Segment End Ratio100%

Segment Reprocessing Thresh0%

Iterative Processing

OCR Iterations1

Enable Cell ValidationFalse

Results Filtering

Minimum Character Confidence0%

Eliminate Edge CharactersFalse

Eliminate Isolated SymbolsFalse

Minimum Character Size

Image Cleanup

Auto Invert ZonesTrue

DeskewTrue

DesshadeTrue

Invert Whole PageFalse

Noise RemovalTrue

Line RemovalTrue

Document Structure

OrientationAuto

Perform SectioningFalse

Processing Options

Merge Broken CharactersTrue

Reject Questionable LinesFalse

Reject Questionable CharactersFalse

Use LexiconTrue

Allowed Characters!"#\$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNPOQRS

OCR Profile

Defines settings which control how OCR is performed.

See Also

[IP Profile](#), [Rectangle](#), [Bounded Region Reader](#), [Logical Size](#)

Used By

[Data Column](#), [Data Field](#), [Data Model](#), [Data Section](#), [Data Table](#), [Grooper Voting](#), [Field Zone](#), [Registration Zone](#), [OCR](#), [PDF Text Extract](#)

Transsym OCR 4

Transsym OCR is a commercial OCR engine which provides highly accurate English-only OCR.

Remarks

- Used for Full Text and Zonal OCR of documents containing machine print characters.
- Does not support handwriting.
- Fully installed with Grooper.
- Requires 1 license per concurrent process.
- English language only.

Change the following property values:

Property	Default Value	New Value
General		
IP Profile	-	OCR Cleanup
Bound Region Processing	Disabled	Enabled
Synthesis Options		
Segment End Ratio	100%	125%
Segment Reprocessing Threshold	0%	90%
Iterative Processing		
OCR Iterations	1	2
Enable Cell Validation	False	True
Rows	-	1
Columns	-	4
Skip First Column	-	True
Results Filtering		
Minimum Character Confidence	0%	20%
Eliminate Isolated Symbols	False	True
Document Structure		
Orientation	Auto	None
Perform Sectioning	False	True
Processing Options		
Reject Questionable Lines	False	True

FileEditToolsHelp

RefreshAddDeleteRenameClone

ACE Architect 02

- Batch Processing
- Content Models
- Data Extraction
- Global Resources
 - IP Profiles
 - Lexicons
 - OCR Profiles
 - Downloads
 - Invoices OCR
 - Scanner Profiles
 - Separation Profiles
- Infrastructure
- Reports

OCR ProfileOCR TestingAdvanced

SaveCancel

General

OCR EngineTranssym OCR 4

IP ProfileOCR Cleanup

Region of Interest(0,0):(0,0)

Bound Region ProcessingEnabled

Description

Synthesis Options

SynthesisEnabled

Font PitchAuto

Maximum Variance12%

Segment End Ratio125%

Segment Reprocessing Thresh90%

Iterative Processing

OCR Iterations2

Enable Cell ValidationTrue

Rows1

Columns4

Cell Edge Buffer0.1

Cell Overlap0.25

Skip First ColumnTrue

Results Filtering

Minimum Character Confidence20%

Eliminate Edge CharactersFalse

Eliminate Isolated SymbolsTrue

Maximum Symbol Ratio1

Image Cleanup

Auto Invert ZonesTrue

DeskewTrue

DeskewTrue

Invert Whole PageFalse

Noise RemovalTrue

Line RemovalTrue

Document Structure

OrientationNone

Perform SectioningTrue

Processing Options

Merge Broken CharactersTrue

Reject Questionable LinesTrue

Reject Questionable CharactersTrue

Use LexiconTrue

Allowed Characters!"#\$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNPOQR

OCR Profile

Defines settings which control how OCR is performed.

See Also

IP Profile, Rectangle, Bounded Region Reader, Logical Size

Used By

Data Column, Data Field, Data Model, Data Section, Data Table, Grooper

Transsym OCR 4


Transsym OCR is a commercial OCR engine which provides highly accurate English-only OCR.

Remarks

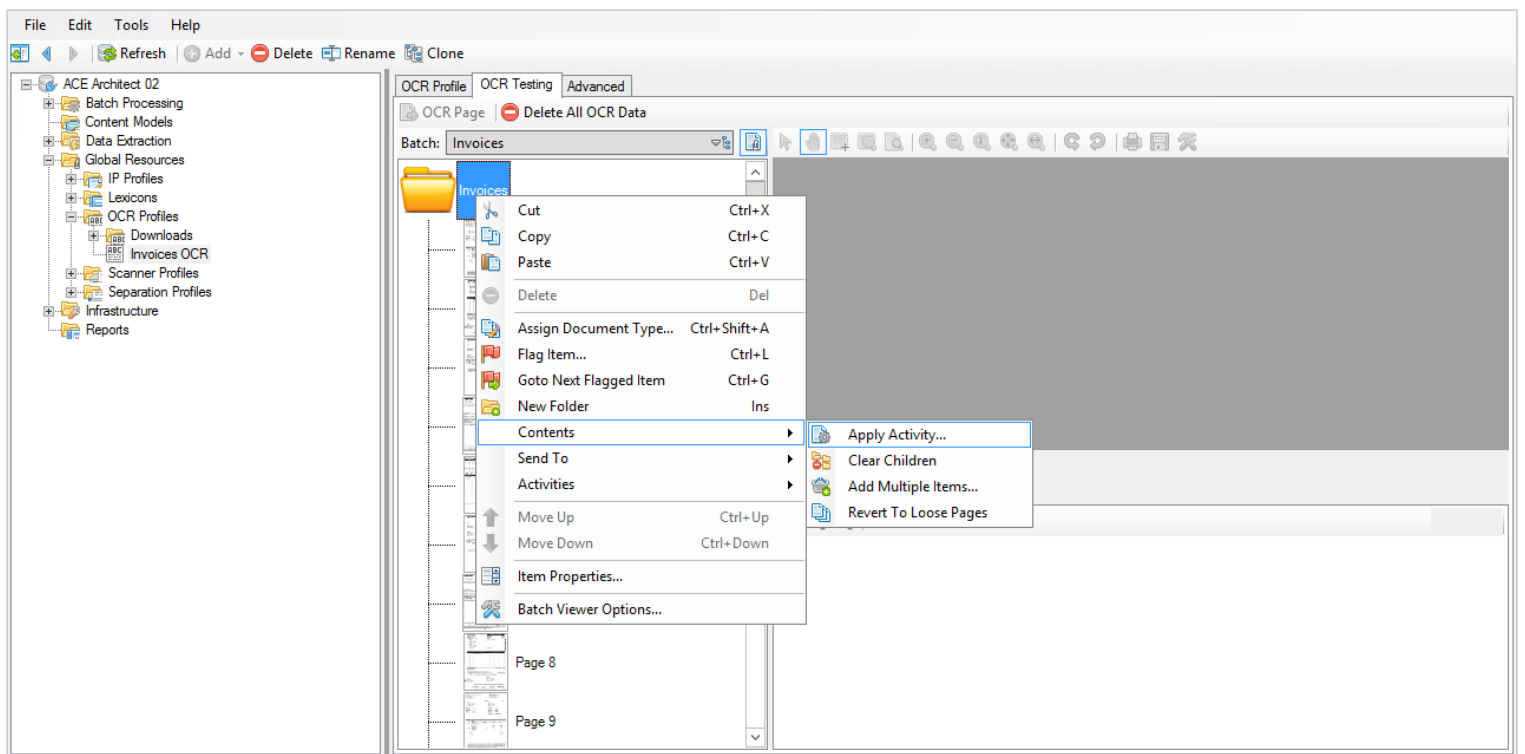
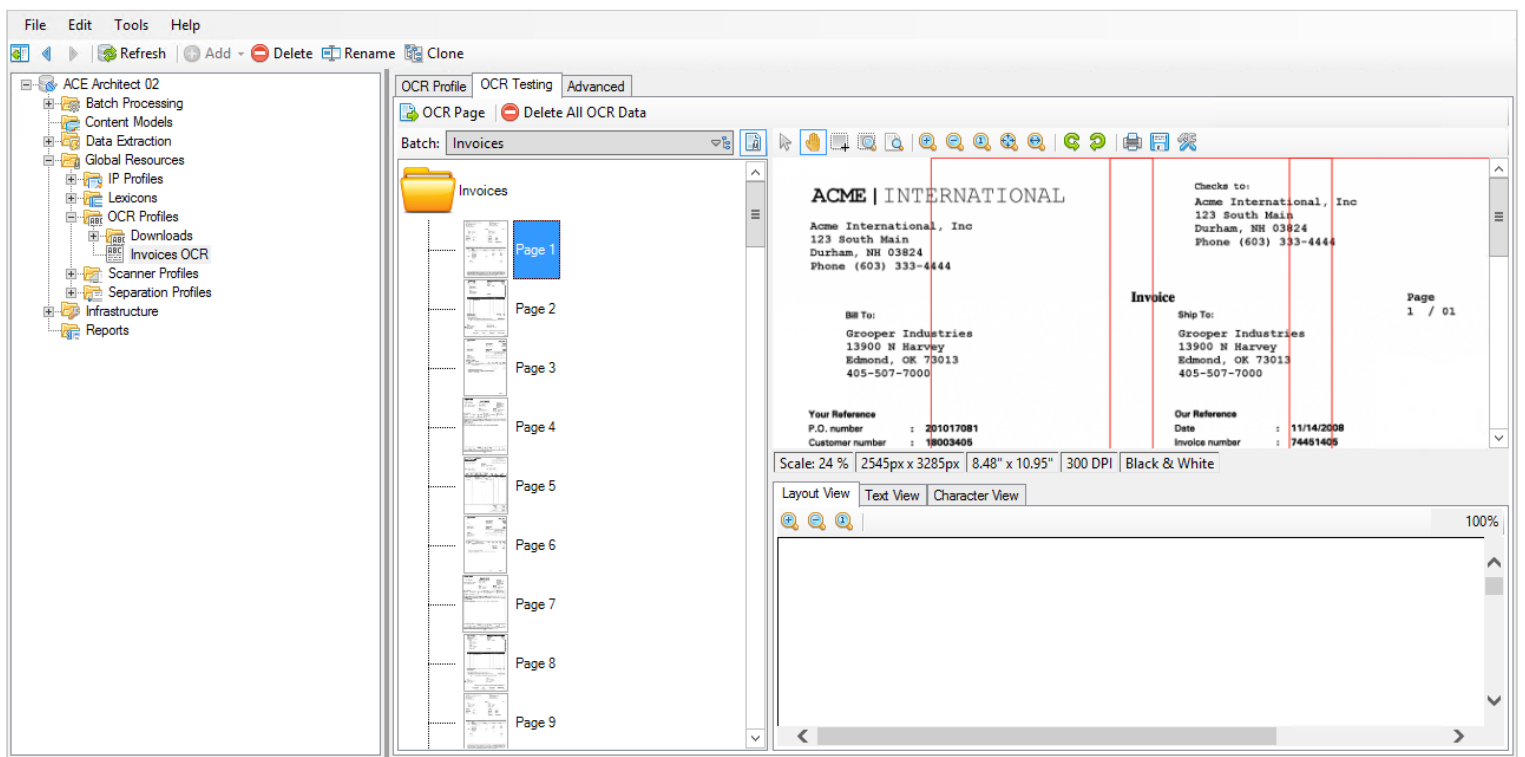
- Used for Full Text and Zonal OCR of documents containing machine print characters.
- Does not support handwriting.
- Fully installed with Grooper.
- Requires 1 license per concurrent process.
- English language only.

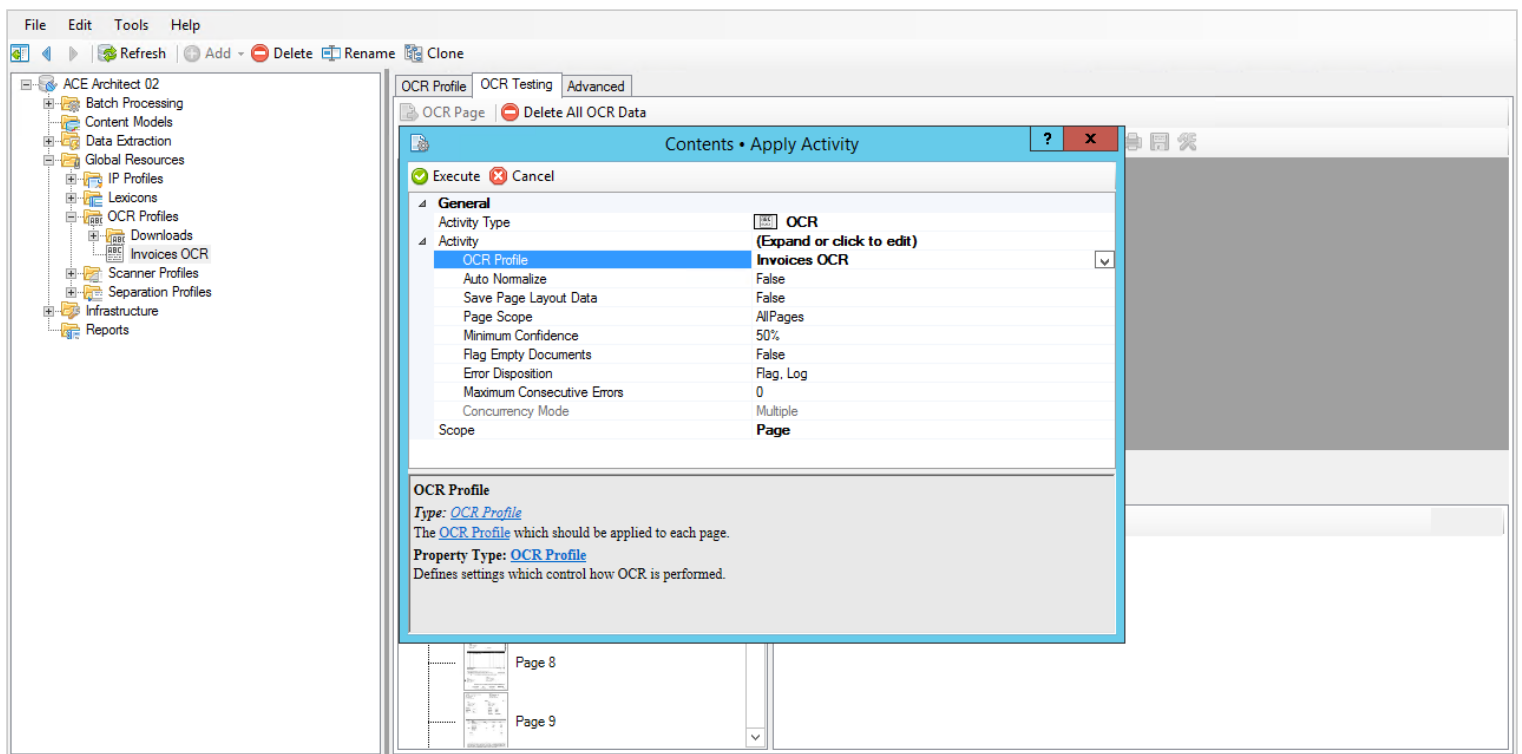
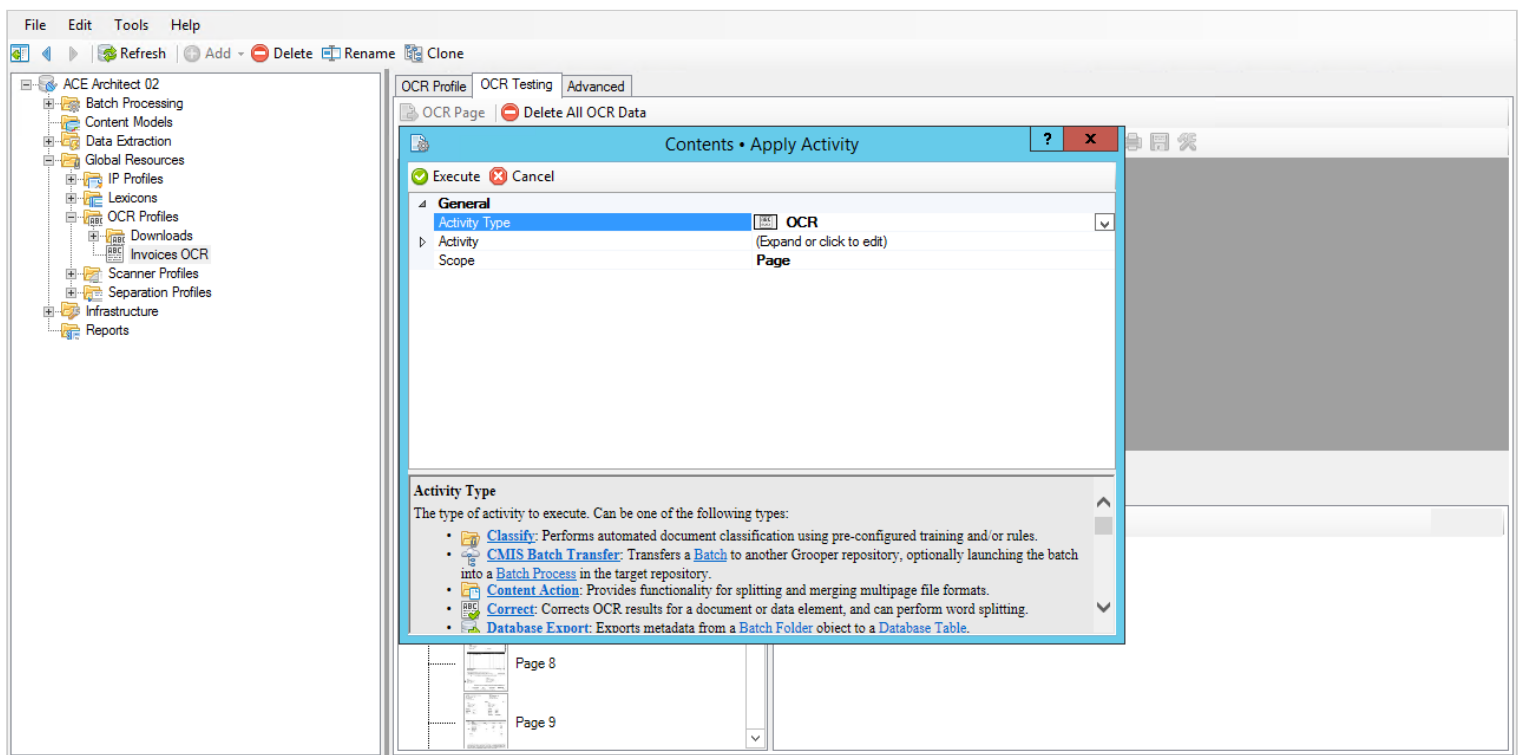
Running OCR against the test batch

- 1. Switch to the OCR Testing tab.
- 2. From the batch dropdown select the Invoices batch.
- 3. Right click on the root batch folder and select Contents > Apply Activity... .
- 4. In the Contents • Apply Activity window, change the following property values:

Property	Default Value	New Value
General		
Activity Type	-	OCR
Activity		
 OCR Profile	-	Invoices OCR

- 5. Press Execute .
- 6. Press Start .
- 7. When the activity has completed, press Close .





FileEditToolsHelp

RefreshAddDeleteRenameClone

ACE Architect 02

- Batch Processing
- Content Models
- Data Extraction
- Global Resources
 - IP Profiles
 - Lexicons
 - OCR Profiles
 - Downloads
 - Invoices OCR
 - Scanner Profiles
 - Separation Profiles
- Infrastructure
- Reports

Process All

StartStopCloseThread Count: 4

General

Batch	Invoices
Activity	OCR
Total Tasks	73
Status	Running
Thread Count	4

Processing

Total Tasks	73
Tasks Completed	9
Errors	0
Tasks / Minute	37.00
Time Elapsed	00:00:14
Estimated Time Remaining	00:01:43

Task Stats

OCR Characters	11,899
----------------	--------

2018-10-17 05:00:03 PM
Processing 73 items on 4 threads.

Page 9